



18 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENT- UND
MARKENAMT

12 Offenlegungsschrift
10 DE 198 46 493 A 1

51 Int. Cl. 7:
C 12 N 15/52
C 12 N 15/63
C 12 N 1/21
C 07 D 493/04

21 Aktenzeichen: 198 46 493.2
22 Anmeldetag: 9. 10. 1998
43 Offenlegungstag: 13. 4. 2000

DE 198 46 493 A 1

71 Anmelder:
Gesellschaft für Biotechnologische Forschung mbH
(GBF), 38124 Braunschweig, DE

74 Vertreter:
Patentanwälte Dr. Boeters, Bauer, Dr. Forstmeyer,
81541 München

72 Erfinder:
Beyer, Stefan, Dr., 38124 Braunschweig, DE; Müller,
Rolf-Joachim, Dr., 38124 Braunschweig, DE

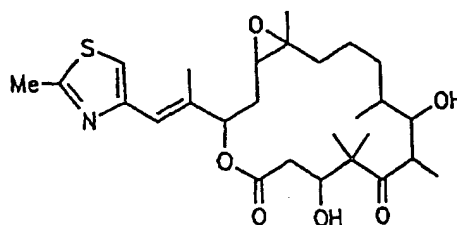
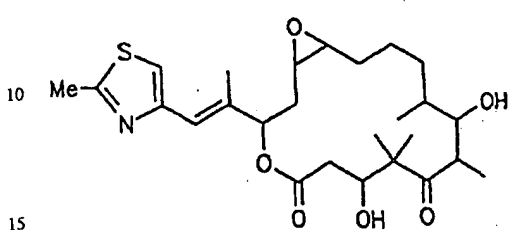
Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

- 54 DNA-Sequenzen für die enzymatische Synthese von Polyketid- oder Heteropolyketidverbindungen
- 57 Die Erfindung betrifft eine DNA-Sequenz, deren Expressionsprodukte eine enzymatische Biosynthese, Mutasyntese oder Partialsynthese von Polyketid- oder Heteropolyketidverbindungen bewirken oder daran beteiligt sind.

DE 198 46 493 A 1

Die Erfindung betrifft DNA-Sequenzen für die enzymatische Synthese von Polyketid- oder Heteropolyketidverbindungen nach Patentanspruch 1, und zwar insbesondere zur enzymatischen Synthese von Epothilonen.

- 5 Polyketid- oder Heteropolyketidverbindungen, insbesondere Epothilone, der folgenden allgemeinen Strukturformel sind beispielsweise aus DE 41 38 042, DE 196 47 580.5 und DE 197 07 501.6 bekannt:



EPOTHILON A

EPOTHILON B

worin R^1 Wasserstoff, C_{1-4} -Alkyl, C_{1-4} -Acyl, Li^+ , K^+ , Na^+ , $\frac{1}{2} Mg^{2+}$ oder $\frac{1}{2} Ca^{2+}$ bedeutet und R^2 Wasserstoff oder eine Methylgruppe darstellt.

- 25 Die Epothilone werden in die Typen A bis F eingeteilt. Sie haben cytotoxische und/oder immunsuppressive sowie antibiotische und antifungale Wirkungen und finden daher zum Beispiel Anwendung als Mittel für den Pflanzenschutz in Landwirtschaft, Forstwirtschaft und/oder im Gartenbau.

Die Epothilone wurden bisher fermentativ durch Kultivierung von Sporangium-Stämmen hergestellt und durch Anwendung herkömmlicher Techniken isoliert und gereinigt, vgl. z. B. DE 41 38 042.8.

- 30 Fermentative Techniken sind aber oft mit Nachteilen verbunden. Der produzierende Mikroorganismus erlaubt nicht in jedem Fall die fermentative Herstellung in großem Maßstab. Häufig kommt es zu Komplikationen bei der großmaßstäblichen Kultivierung oder die Ausbeuten sind gering oder die Isolierung und Reinigung sind aufwendig.

- 35 Daher wäre es vorteilhaft, wenn zur fermentativen Herstellung der gewünschten Verbindungen ein gut charakterisierter und leicht zu handhabender Mikroorganismus zur Verfügung stünde. Wenn ein solcher aber nicht in der Natur gefunden oder gezüchtet werden kann, bleibt nur noch die entsprechende Veränderung eines geeigneten Mikroorganismus mit gentechnischen Methoden. Dazu ist aber die Isolierung und Charakterisierung der entsprechenden Gene erforderlich.

Aufgabe der Erfindung ist daher gemäß Patentanspruch 1 die Bereitstellung einer DNA-Sequenz, deren Expressionsprodukte die enzymatische Biosynthese, Mutasyntese oder Partialsynthese von Polyketid- oder Heteropolyketidverbindungen bewirken oder daran beteiligt sind.

- 40 Durch die Bereitstellung einer derartigen DNA-Sequenz lassen sich folgende Vorteile erzielen.

Die DNA-Sequenz läßt sich mit üblichen molekularbiologischen Methoden in bekannte und optimierte Expressionsvektoren inserieren, wodurch die entsprechende Transformation, Selektion und Klonierung von Zellen möglich ist, die dann zur Synthese von Polyketid- oder Heteropolyketidverbindungen durch Fermentation in der Lage sind. Wenn ein überproduzierender Zellklon gewählt wird, lassen sich die gewünschten Polyketid- oder Heteropolyketidverbindungen

- 45 leicht in großen Mengen herstellen und gewinnen.

Die Kenntnis der Lage der regulatorischen DNA-Abschnitte und der einzelnen Strukturgene gestattet die gezielte Mutagenese ("site-directed mutagenesis") mit üblichen gentechnischen Methoden und somit die Konstruktion von optimierten Enzymen ("protein engineering") zur fermentativen Synthese von Polyketid- oder Heteropolyketidverbindungen.

- 50 Die Erfindung betrifft somit ferner einen rekombinierten Expressionsvektor nach Patentanspruch 8, damit transformierte Zellen nach Patentanspruch 9 sowie ein Verfahren zur enzymatischen Biosynthese, Mutasyntese oder Partialsynthese von Polyketid- oder Heteropolyketidverbindungen nach Patentanspruch 15.

Vorteilhafte Ausführungsformen der Erfindung sind Gegenstand der Unteransprüche.

Die Erfindung wird nachstehend detaillierter erläutert.

- 55 Fig. 1 ist eine Restriktionskarte der erfindungsgemäßen DNA-Sequenz, die auch die Lage der regulatorischen DNA-Abschnitte und der einzelnen Strukturgene ("open reading frames" (ORF) 1 bis 14) angibt.

Fig. 2 ordnet den ORF 1 bis 14 die jeweilige biologischen Funktion (Regulatoren, Enzyme) zu.

Isolation und Charakterisierung der DNA-Sequenz

- 60 Es wurde genomische DNA aus dem Myxobakterium *Sorangium cellulosum* Soce90, Stamm und Anzucht bekannt aus DE 41 38 042, verwendet.

- 65 Genomische DNA wurde mit Hilfe des Qiagen Blood & Cell Culture DNA Kits (Qiagen, Hilden, FRG) isoliert. Siehe dazu "Genomic DNA handbook" S. 31 ff (Qiagen 1995). Modifizierungen: Nach Denaturierung und Proteolyse wurde eine Phenol-Chloroformextraktion, gefolgt von einer Ethanolpräzipitation, vorgenommen (Sambrook J., Fritsch E. F., Maniatis T., Molecular Cloning, A Laboratory Manual, Cold Spring Harbour Laboratory Press, New York; 1989). Die in Puffer gelöste DNA wurde anschließend auf die Qiagen-Reinigungssäulen aufgetragen und gemäß den Herstellerangaben isoliert.

Konstruktion von geeigneten rekombinierten Expressionsvektoren

Expression in Myxobakterien

Eine heterologe Expression der in Fig. 1 aufgeführten ORFs wird unter Verwendung eines Derivats des Plasmids pSUP102 (Simon R., Prierer U., Pühler A.; Methods in Enzymol, 118: 643-659; 1986), bei welchem das Chloramphenicolresistenzgen durch eine Streptomycinresistenzgenkassette mit Promotorelement aus dem Transposon TNS ausgetauscht wurde, ausgeführt. Homologe kurze Abschnitte genomischer DNA aus dem Wirtsorganismus werden mit den DNA-Sequenzen entsprechend Fig. 1 unter Benutzung effektiver Regulationselemente in z. B. die Restriktionsschnittstelle EcoRI des Vektors ligiert. Nach Amplifikation des Vektors in Escherichia coli erfolgt der Transfer der DNA durch Elektroporation der Wirtszellen oder durch Konjugation mit Escherichia coli S17-I (Simon R., Prierer U., Pühler A., Bio/Technology 1: 784-791; 1983).

Mit Hilfe der durch den Vektor vermittelten Tetrazyklin- bzw. Streptomycinresistenz werden die Wirtszellen auf Integration der rekombinanten Plasmid-DNA durch homologe Rekombination in das Chromosom überprüft.

Expression in Zellen von Streptomyces

Eine heterologe Expression der in Fig. 1 aufgeführten ORFs wird unter Verwendung der bifunktionalen Streptomyces-Escherichia coli-Cosmide pKU206 oder pOJ466 vorgenommen.

Expression von Zellen in Escherichia coli

Eine heterologe Expression der in Fig. 1 aufgeführten ORFs wird unter Verwendung von "Bacterial Artificial Chromosomes", Cosmiden (z. B. Supercos; Stratagene GmbH, Heidelberg) und T7-Expressionssystemen (Stratagene GmbH, Heidelberg; New England Biolabs GmbH, Schwalbach, FRG) vorgenommen. Die Expression rekombinierter Enzyme erfolgt in Escherichia coli-Zellen, die eine konstitutive Expression einer Phosphopantetheinyl-Transferase gewährleisten, welche für die Bildung von Holoenzym-Polyketidsynthasen und -Polypeptidsynthetasen notwendig ist.

	CGATCGCGGCGCCCTCGCGTGTCTCTCGAGCGTGCAGGCACTCCACGCCAGCGCGACTTGGCGAGGCCAGGCGCGCCACCACCA	50
	CCACCGCTTCGCGGAGGGCTCGTCGACGCAATGGCGCACTCGGTGCGAGCTGCGAGAGCTCGCGCTCCCGCCCCACGAGGGCTCG	130
5	GCTTGGCGAGGAGCGTGGCAGGCATCGGCTCTCTCTCGGGCCGCAAGCCAGCACCTTCCGGCCCCGTGTACGCTTCGAAGCGGC	270
	TGCGGAGCAGGCTGCGCGTCTGCTCGTGCAGCGGATCTCGGGCGGACAGGCCATCTCGCCCGGCGATGAGCTGCGCGACCCGATCGA	350
	CCAGCTCGCGACCGCGAGCTCGCTCGACCTCGGCGAGCTCTGTCGCGACGGACACGGGACGCTTCGAGCGCGCGCCGACGCGCA	430
10	CGCGCGAGTGGGCCCGCGGTGGCGAGTCCGTGGGCGACTCGGCGCGGACACCGCGAGGCCACCAGCGCGCTTGCAGCCCATCGA	510
	GCGCGCGCGCTGGCGCGCGCGATGTCCGCGAGCGCTCGGGCGCGCGCGCGCTCGTCTCGGAGAGCTGCGCGCGCGCTCGCGCG	590
	CGCGCTTTCGGCGAGGATGACGACATACCTTGGCTCGGCGCTCGTGTGCTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCTCG	670
15	CCCCGATCGAGAGCCCTTCGCGCGCGACGGCGCGAGCTCGCGCGCGCGCGCGCGCTCGCGCGCGCGCTTCGCGCGCTTCTTCGCA	750
	GCACTCGCGCGCGCGCGCGCTCG	830
	GGAGCGCGAGCGCGCTGTTCGCGAGGAGCG	910
20	CGTCGACCGCGCGCTCG	990
	GCGCGAGTGGCGCGCGATGCGGAGTGCAGCGAGCTGACCGCTCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1070
	GGTGAACG	1150
25	CGCGGAGCGCGCTTCGCGAGGCTGCG	1230
	GCACCG	1310
	TGACCG	1390
30	GATCTCGATCACATCGCTGCG	1470
	GCTTGAACG	1550
35	AACCG	1630
	TCACACG	1710
40	AGCGCTACTTCTGCTGCG	1790
	ACGTGCGCGCTCTGCTGCG	1870
	GCACCG	1950
45	CCTGGGCTACG	2030
	AGATCCG	2110
	CGTGGCTCGACAGCG	2190
50	TCGCAACACTCTACGCTCAGCACCG	2270
	AGCGGTGGCTCGCAACGCGCTCG	2350
55		
60		
65		

GCGCCAAGBAATCTCATCGAGCGGCAGGACACACCTCGACAGCCTGGTGGATCGCCTGCCGAGCCGCGCATCCGCGCGTGTG
 2700
 AGCCGATGCTCGCCGGCACCCGCTTCCGAGCGTGCCCCCGACACCTTCGTTTCGCGATCGACCTCCGCGCTCGTGCGCATGACCGCG
 2790
 AGGGCGGCTCGACGTGCGCAACCCCATCTACCGCGAGATCATCGTCCGCGAGCTCGCGTTCCCGATCCGCGCCTCACTCCCCAGATCA
 2880
 AGGCCACGTGGCTACGCGAGGACGGCGGCTCGACCGGACCGCTTCTGACGCTTCTCTCTCTTCTGCGCGGACGCGCGAGCCCG
 2970
 TCTCGCGCGCGCGCTTACCATTGAGATCGCCCGCACCTCGTGGTGTATGCGCTTCTTCCACCGCGTGTGAACGCGCGTCCGACCGT
 3060
 AGCGCGAGTACGCGCATCGCGCGGGCAGGATGGATCTCTGCGTTCTGTTACGCGGCGAGACGCTCGCGATCGACCTCAAGGCTGCGGAG
 3150
 ACGGCGCGCGCGATCCGCTCGCGAGGGCTCGCCAGCTCGACGAGTACCTGCGCGCGCTGGCGCTCGATCGCGGATGGCTCATCTCT
 3240
 Eco47 III
 TCGACACGCGTCCGACAGCGCCCCATCGCGAGCGCACGCGCGCGAGCGCGCTCTCCCGCGCGCGCGAGGTCGCGCTCATTC
 3330
 GCGCTCGAGGAGCTCGCGCGCGCGGCGAGCGCTTCCACGAGCGCGCGGCGACCTCGGACAGCGTCTCTACTCTCCGAGCGCGCGCG
 3420
 CCCCCCGCGCGCGCGCGCGCGCGCGCGCTCCAGCTCCAGCGCACCAAGGACCTCGCGATCCGCGAGCTCGCGCGCTCCAGCA
 3510
 GCGCGTGGCGCGCGCGCGCGCGCGCGCTCCAGCGCGCGCTTCCAGCGCGCGCGCGCTCGCGCGCTCGCGCGCTCGCGCGCTCGCGCG
 3600
 GGTGCTCAGCAGCGCGCGCGCGCGCGCGCTCGCGCGCGCGCTTGGCGCTTGTGACCGCGCGCATCGCGCGCGCGCGCGCGCGCGCG
 3690
 BamHI
 CCGGATCCGAAGCG
 3780
 ACACCCAGCG
 3870
 CGCGCGAGCG
 3960
 AGAAGAAATCTCGGTGCGTCTGAGCG
 4050
 Bgl II
 GCAAGAGCGCGCGCGTCTGAGCTCTCGGAGATCGGGTGGACCTCCGCGCGCTGCGTGAACAGTACTTGTCTCGGTTCAGATCTTCTGTA
 4140
 CGAGCGCTTCTGCGATCTTCTGAGCCTTCTGCTCGAAGCGCGTGTCTGCGCGAGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCG
 4230
 BamHI
 AAGAATACGTGTCTGAGCAGGTGCTATCGCGGTGACGCGTTCGCTTCTCTGAGATCTTCTTGGATCCGCGATCGAGGATCCACCGCG
 4320
 AGATCGCGAGCTGTGCG
 4410
 CCTGCGCGCGCGATCG
 4500
 TGAACACGCTCTGAGACGCTCGGACTCGCGCGTGAAGCG
 4590
 TCGGTCGACCGCGGAGCATCTCGCGCGTCTCGGAGCG
 4680
 ACTGACCGCGAAATACCGTGGCGGCTGATGCAACGCTACCTGGAGCG
 4770
 GGTGCCACTTGTATCTTGTGCGCGTATTCGAGGAGCTCGCGCTTCTTGTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
 4860
 CGCGCTGGTGGCG
 4950
 CGTCTCTCTCGCGCGGAGCGCTCGACCG
 5040

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

5130
 5220
 5310
 5400
 5490
 5580
 5670
 5760
 5850
 5940
 6030
 6120
 6210
 6300
 6390
 6480
 6570
 6660
 6750
 6840
 6930
 7020
 7110
 7200
 7290
 7380
 7470

Eco47 III
 BamH I
 BamH I
 BamH I

GGGCCGGGTCGCGGCTCTCCACGCGCCCTCGCCGAACGTACCGGACAGGACACGGCCGTTCTTGCCGAGCATCTCCCGAGCGGAGCT
 TCTGCTCCCGCCACCAGATCAGTCTGGTCGGTCGCGGAGGTACAGACCATCAGGATGCCCGTGCCTTCTCGCGATCCACGAGCGCGC
 TCGGGAAGATCCGACCGCGCGCGGAGATCGGGTGAGCGCCGCTTGCCGAACAGGTGCTGATACCGCGGGTCTCCCGGTGCGCGC
 TGACGCGGACGAGGCGCGGAGCAGCTCCGGGCGGCTCGTGGCGATGACGAGCTCCTCGGCCGTCCCTCCACGCGGAACGCGATGTCT
 GGAACGCGCCGATTCGCGGCGATCCTCGACCTCCGCTGGGCGACCGCGGTCTGGAATCGACGTCCACATCGTCCCGCGGAGACCG
 AGTAGACCTCGCCCTTCTCGTGGAGATCCAGGAACGACAGCTGCGCGCTCTGCGGCGATGATCATGATGGTGGCGTACTCGTTCGCGC
 ACTCGACCGAGAGGCCACCGCGGGAAGAGCGCTTGAGACCTGCTCGTCTCGCGCGTGACCTTGCGGCAGAGCTCGATGAAGTGG
 GCGCGACACGATGCGCGCGGCTCTCTTGTATCGTCTCGCGCGCGGCTGCGGCAAGGTACGGCCGCGCTCGTACGCGGTGCGCACGT
 CGGTGCGGACGTGAAGTAGTTCTGCACGCGCGCTCGGTGGGACGCGGTGCTCGTCCACGCCATCGGGTAGAAGATGTTGAAGCGC
 GCATCCGGCGTGGCGGACGACGACGCTCGTGTGCTGATGCTGAAGAGTGGCGGATGTGCAAGCGAGCCGAGGCGGTGCGCGCGCGG
 TGTGACGACGAAGGTCTCTCGCGGGGCGCGACGGGTGCTATCGTACGTCCCGTGGGCTCCACAGGTGCGCCAGGCGAGCTCGG
 CGACCGGCGAGTGAAGTCTCTCGGAGCGCTCGCGGATCGATCGAGCGGAACGCTTCTTGTATCGTACGTGCTACCTGACGAACAGA
 CCGCGAGGAACCGCGCGGGGCGGATCTCTACGTCTCGCGCGGCTGCGCTAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
 ATCGCATCGCGCATCCCGCAGAGCGCGCGGCTCGCGCGCGCGCGCGCGCGCTCGGTGGAGCGGAGAGAGCGCGCGCGCGCGCGCGCT
 GTGGAGGACCG
 CGTGGACGTGCTTGGGTGGAGCGCGAGCTCGCGCGGGCGAGCTCGCGGACCGACGAGACCATCCGATCCAGGCCGAGATGAACACGT
 GCGCGCGCGGATCTCCGCTCTTCTCGCGAGCTCCCGGTGAGGCTCGGGCACGTGCGGCTGCGAGTACCGCGCGCGCGCGCGCGCGCGCG
 GCGCGCGCGCGGAGCGGTGATCTCGTAGCGGATCCGGTCGGATCGCGCGCGGAGCGCGCTCGAGGTGCTCGCGGTAGATGAGCTCTCT
 CGAAGCGCGCGCGGAACAGGATCCACAGTGGGGCG
 CGGTGCGCGTCTCGGACGAACAAGGAGGGCGCGGAATCCCGGATCGCGGTGAAGAGCGCGGTGCGGGCGGATGGCGCGGAGCGTGGCGC
 CGGGCTCGAGCGGTCAGGTGCTCGAGCG
 ATCGGATGAGTAGGCGCGCTTACCTCGCGCGCGGGAGCGGGAGGAGGAGGTGACCCATGGCGCGCGCTCGAAGAGAAACGACCTCC
 CCGCGCGCGCTCGAAGAGAGCTCGCGCACGAAGGGCTGAGGGCG
 TGGTGGGGCGCGCGCGCGCGCTCGGCTGCGAGGCGCGCGCGGTGGCGAGGTETACCGCAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG
 GGGAGGGCGCGCGCGGAGGCGCGCAATCGCGGAATACGGGCTTCCGCGCGTGGCGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG
 GCGCTCGGGCGGGAGCGCTCGGGGCTETACCGCAGCGCGATCGCGCGCGCGTGGCGGATGTTACCGCGCGCGCGCGCGCGCGCGCGCGCG
 TGGGGAGCGCGCGCGCGCGGAGTCCGGTGGCGCGGCTTCTGCGCGGGGAGCTGCCCTCGCTGAGCGCGCGCGCTGGTTCGGCGGACT

TGGCCGAGCCAGCGCGAGATCGGAGACGGGGAGATGCGCGCTCGCTGCATAGAATCTCCATCGAATCGGTATCAACACATCGGCAA
 7590
 GAGCAGCCAGGCTCAAGAAACCTTCGAAGAACCGGCTCTCATACACCTCCATTTCATCGTCCGACCCCGGATTTCAGGACGGATCGAACC
 7650
 CGCGAGGACGCTGGCTCTCTGGGCTCTCCCTGCTCGCTCGACCGCGCCCTCTCGACGCAACTCCGCCGTTCGTCGGACGGGACGGT
 7740
 CCGCTCGCGCAGCTCCCGTCGAGACGACTCAGCGTCTCGACGTGAGGAGAGATGACGACTCGGCCCTCGCGCCACGACCTTCGG
 7830
 GCTCGGTGCTTCGAGCGCGCGCCAGCGAGCGAGCGGCGATCGCCAGGAGATCACGAATCTCCCGCCATCGGCCCTCAGCGCTCGGGC
 7920
 TCGTTGCTGCTCGTCCCGCTCGCTCCCGCTCGCGCGACGCGAGCTCGCGCGGGGAACCGCGGGCGCTGCTGCTGCTGCTGAAGCG
 8010
 CGACGATACAGGGGACCGCTGCTACCTGGGCAACAGGCGCTCATCTTCTAACCAGGCGAGCACTACGGTGAGTGTGCCATGAGTA
 8100
 GGGCCCTGAGGGTCCGCGCGAGGAGCGTGGTGTACGCGAGAGATGCGCATGGTGGACCGGGCTACGCGTCGAGAGGGGACACTAGCACT
 8190
 CGACCTCGATCTGCGCCAGCACITTTTGTGGGAGGGGCTGCCCTCCCGCTGGCGCTGGCGCTGGCGCTCGCGCTGGCGCTGGCGCT
 8280
 GCTGGCGCTGCGCGCTGCGCGCTGGCGCGCTGGCGCTGCTGCGACGTGAGCTCGAGCAGCCCGGGCTGACGGACAGACCCCGAGTTCAT
 8370
 CGAGCGGTGATGCCGAACCGCCAAAGCGAAAAACGTATCCGTTCCGAGGTCGTGGCTATCATGAAGCTGCTCGATGCGCTGACAG
 8460
 GCTTCTCGAGATCTCTGCTGCTTTTGGCAAGCAAAACCGCATGAAGCGACTCCCTTGGCTCCCTTCAAGCAAGCGCTCGCTGCGCAG
 8550
 CCGCCACCCCGGCTCGTCGAGCAAGTAGATGGCTCGCTGCTGACCTGCTCTCCCGGTAGGCGAGACACATCCCGCAGCAGTAGTAC
 8640
 GTCCCTGCGGCGCGAGGCTGGCAAGCGCTTTCTCCAGCGCGCGACAGAACCGGTCTCGCTTCGTTTCATATCCCTGGGCAAGCCCC
 8730
 BamH I
 GTGTAAACGAGCGAGGAAGGCGCGGATCCCGGACGCTCCATGCTGACGCGGCTCGCGCCCGACATACAGCAGGTGCTCTATG
 8820
 GCTCCAATGGCTTCGCGCCACCTGGCATCGGCGACGCTGATGCGCATCCGCTATCTGTGATGCTGAAGGCTTCGAGTAGCCGCTATC
 8910
 GTGATCGTACGCTCGACATCGCGGAAGGGAGGCGACCGTGACGTGCTACGGCGCTCGAAGATAAAGTACTCGTAATTTCTGCTCGT
 9000
 ATCACCATGAGGTGATGGTGGCAGCGAGATCGCGCATCTGTTCCAGCTCCATTCGCGCAACACCTTCCCGGAAGGATTCAGGAGAG
 9090
 TTCAACGATCGCTTGGTCTTCGGGCTGATCGCGCGCTCCAGCTCGTCCCGCTCGACATTCCAGCTCAGGGATCGCGCGCTCACATAC
 9180
 CGCGGAACAGCTCGACGCGGAGGATAGCCTGGCGGTGATAGGCATAAAACGCTCGAAGAGCAGCACTTCGTCCCCAGGATTGAGCAAG
 9270
 BamH I
 GCCATGCAAGTGGCTGAAGGCGCTGCTGCTCCGCGCTCACCGTATGTCAGTCTCCGATCCCGCGGATGCCATTATGGCGAGCC
 9360
 AGCTTCGCGCGATCGCATGGCGCAGCTCCACGATGCCGTGGAAGCGCAATATGATTCGACCCCGATCCATCGCTCTCTTACCGCT
 9450
 TGAAGGATCAGCGAAGAACTGGGATACAGAGCGCCCTGGGACATATGATCCATGGACCTTGGCGCACGCGAGGTCATGGTACGG
 9540
 ATATCGGACTGGCGAGGCGAGCGCAGGATCACTGGTAGACTCTTCATCAGCGTCTCTGCTTCTGTTCTGCGGCTCTGCATGGTGT
 9630
 CTTCGGGTGGCTTGTGAGCTCGACGCGCCATGCAAGCGCGCAGCCCTACCGCCGCGAGGCTGTCCACACTTCTTTGATGAAGCGAG
 9720
 ACATTCGGCTTTCGTGCTCTGTTGCGCGCAGCCCTCCAGCCCCAGGTACGGGCTTGTGGCGGGCCAGATCGAGTATGCTCTTCGCT
 9810
 Eco47 III
 GTTCAACGACCTGGCAACGCTCTTCTGCTGCTCCGATTCATGATTTCTCTGCGCTTCGTACGCGTGGCGGAGCATGAAGC
 9900
 AATCGCTCATCGCGCAGCGCGCGCGGCTGCGCGAGGCACTCCACGCTCTCTACGCGAACCTCATCGCTCGGATGTTCCCG
 9990

5

10

15

20

25

30

35

40

45

50

55

60

65

ATGCCGACATCGGATCGTGTAACCTGCCCTGGAGACGGGATATGGGACAAAAATACCTTGCCGCACTTGTGACCTCGTCCACGACGGCCCTCG 10080
TTGAGGCGATCGAGCTCGCGTTTCGATCCACTCTCTCTCTGCCCTCGTCCGCCACCGCATGATGCAAGCGAGCGCCGAGGGCTCATGCGA 10170
5 AAGCAGACCGTACTGAACCGCGTGGCGCGAGGGCGCTCCCAATCGGGATCGGCGTCCACCCACTCGGCGAGCTGCTGCCCAATCGGACG 10260
BamH I Eco47 III
TGCTCCCGATCCGGGCGCCAGCCCTTCATGCCCGAAGTAGCGCACGATCATCCAGAGCTTCAGCGCTCGGAAGCGCCGACCGAGCTGG 10350
10 ATACCCCATGTCATCTAATTCTGTGACGTGCCCCTCGGTCCCGAGGTATTTCGGGCACCAAGACTCAACCGCGCTTCAGTCGGTGGGCTCA 10440
CGCAGTAGAGACCGCTGCAATCCATGGGGTGAAACAGCCACTGTGACGGTTCACTACCAGGAGTCCGCCCTTCGCAACCGCGGAGC 10530
ACGTCCCTGTGCTCGGGGACGATCGCGCCATCCCGCGGTAGGCGCGCTCACGTGAAGCCATAGCCCGTGTCTCCGGCAAACTGATGAG 10620
15 ATGGCGGGGATGGGGTCGACGCTCGTCTGGACGTGCTGCCCACTGCGCGGACGACGAGAAGGGTCGGAGCGCGCCCTCGAGGTCTCTC 10710
BamH I
ACGACGCGCGCGCGCAGCGCTCGGGGACCATGCGGAAGGCGGATCCGTGGGATCTTCGCAACCCCTCTGCCCGATGCCAGGGTG 10800
20 ATGGCTGCCCTTCGATGGATGATGCGCTGTCTCGACGCGTAGATGCGCATCGCGCGCTGTCCCGCCATGCCCGGAGCGGATGCTC 10890
GGCTCGCGGAGTCCGCGCGGCGCGGATCGCGACCATGCTGGCGGTGACGCGGTGTCCATGATCGCGCGCTGCAAGCGCGGCTGAGAA 10980
TCCAGCATCTGACGCGAGCGAGAGGACGAGCTCTCGAGCTCGGTGGCGCGCGGCGAGCTGCCCATAGCATCAGTTGACGTGAGG 11070
25 BamH I
CACGCGCGGAGCAGCTCGCGGAGGATCCAGGACACAGCGCGTGTTCGCGAAATACGCGAAGATCGCGGATGATCCAGTGGGTGATC 11160
CCCGCGAAGATGATCTGCTCGAATCGGTGAGCACGGGCTCCATCGGCTCGGCTGACGCGCGGGTCCGGGCCAGCTGCCCTTCAGC 11250
30 TCGCGGGGCGGATCGCGGAAAGACGGGCTATCGATCGGGTGCGCGAGTAATCGCGCGCCCAATCGATGATTCTCATACCGATCCG 11340
CGGAATCTCTCAGATCCATGTCCCGAGCGGTTCTTTCCGCGGGTGGCTCACGTCAACCTCTCGGCCCTGCCAGGACAGGATCTCGAG 11430
35 Eco47 III
GTCCCTCGGCTCGGCGGTGGAAGCGCTCTTTGAACGTGAAGGCCACGGGTCGTGCTGATGCGCGCGAGGTGCTCGAGCGCATCTG 11520
CCCCCTCGCGGACGGACGGGATGTGCCCGGCGGGACCCACCACAGCACAGGTAAATGCGGCTCGAGATGCTCGAACCACCGAGCGCTG 11610
40 TCGCAGGAACCGGATGATCGCGGTGTAGCTGAAGGCGAAAGGCTCGATCGAGGTTCAATACCGACAGGTCACAGGAGCGGCTG 11700
GTCCGGGTACGAGCGGATGGACACAGAGTTCCCTCGGCGCTGACAGCGCCACACGAACCTCTGCTCGGATCGGCGAGATGGTTGAT 11790
BamH I
ATGGTCGAGCCCCGGACGAAGCCCTCCATGATCGGATCTCCAGCGGACGCGAATACATGCGAAGTTGTATTGCGCGTGTGGTGGCG 11880
Eco47 III
ATGCTCCGACATGTGCTTTCCATCTCCAGCTCCCGCTCACCATCCAGCGCTGCTCGGGGAGCTCATCAGGCGACGCGACATCGA 11970
TCCCGAAGCTCCGCGCATCCCTCGACGAAGGCGGCTGGACCGCTTCGCGACGGATCGGCTGCTCGCGCAAGACCTCGGACGACAA 12060
50 AGAAGAACCCTCGTGGAGGGACAATCTTGCCCGCTCGGCTGGCGCCATACGAAGTGGCTGCTGACCACTCCCTCGGCTCGGAT 12150

Bgl II

ACCCGACCTGCGCGGCGCGACCGCCACGCTGCCBCTGAGCCGACCTCCACGAACGCTCACCAGCTCGCGAGATCTCGAGCGCAACGT 12240
 CCGGGCCAGCCAGATCGCGAGGTCCCAAGCGCCGACGGGACGGCGAACCAGCGAGGACAGGAGGTGTAAATCGACCAATCGCTTGA 12330
 TGTGCGGACGTCTCCACACCGAGGACCGCTTCCCGAGCGCTCGATCGAGCTCGGAATTTCTTGGCAGAGACCCCACTCGCTTCA 12420
 TCGCTCGCGCCAGGCAAGCCAGTCCGATCGGACTGGCGTTCCTGCGCGCCAGCTCCGTGCGAGCTCTCTCGACCTTCCGAGGT 12510
 CCTCCAGCAGCGCCGCGCTCTGCGCGTGTGTCAGGCTTCCCGTACCAGGTGACCAAGATCATCCAGGAACGACTCCAGATTTC 12590
 GCGGATCGAGGATGAATGCCATGTGCTCTTCCCTCGAGAGCGATCGCTCGATCGACACAGGCTGTGGATGCATGCGCGGCGGT 12680
 GCGGACGAGGACCCGCTACTCATGGCTCTGCTGCCGATGAACAGGTCTTCCACCGCTCGATCGTCTGGTGGTGGTGGTGGTGGTGGT 12780
 GTCCACCCCGCGCGGACCGGATGTTTCCACCGAGAGCTACGAGGCGCGGTGCGAGCGAGGAACGCCACAGCATCTGCGACCTCGCT 12870
 GCGCGCGCCAGGCGGCGCATGGGAGCGCGCTTCCATCCATCTCTTCTGCGCGGCGCGAGGTATCGCTTGTGATGAGCCCTGCGAGA 12960
 CACACAGTTGACAGGATTTCATGAGCGCTCTCTGCTGGCAGGCTGCGGTGAGGATGAGCAGCGCGTCTTCCGATCGAGTACGC 13050
 CCGCAGCTTCGCGCGCGCGGATCGGTAAGTGGGCTCAACCCGATATTGATGATCCGCGCGCTCTTCTGCTGGCGCATGCGCGCGC 13140
 GCGCGCGGACAGAGTAATGAACGCTGCTAGGTTGCTGTCCATGACGTTGCGCATTCGCTGCTGGTATCGCGCGAGCGGCTTGA 13230
 GAAGAAGTCCCGACCTTATTACGAGGATGTCATGGGCGCCAGCTGCGCTCGAGCTGGAGAAGAGCTCCCGCGCGCTTGGGCG 13320
 GGTGACGTCGCGCTGACACCATGGTTGCTGCGCGGAGCGCGGATCTGCGCGCGCTGCTGCGCGCGCATCTTGTTCGAATGGTA 13410
 ATTGACGCGACGCTCGCGCTTGTCTCGGAGGCGGAGCGGATCGCTTGCATTTCCGCGCGAGCTACCGGTGACGAGGCGGAGCG 13500
 CCGCGGAGCTCCAGCGATCGCGCTGTGCGAGGCGCGGAGCGGCTCTGCTGGAGCTCGAGCTGAGCGGAGCTCCACCTGGTAGCT 13590
 CGTCTCTCGCGGAGCGCGGAGTACCTCTGCTAGAACGCTCGAGGAGCGGCTCTTGGCTGCGATGATGCTCGCGGAGCTTTCGCGCT 13680
 GCGCCAGGATAGAGCAGGATCTGCTGCGGCGCGGAGCTCTGCTGAAAGATCGCGCATGACCGCAGCAGCGCGGCTTCTGCTGGC 13770
 TCTGCGCGGAGCAGATCATCAATTTTGCATGCTTGGTGGCTCGCGCTCCATGCGCGGCTTGTATGCGCATCGCTCCATAACGAG 13860
 GATCATGCTTGGCTCTGTTCTGCTATCGCGGTTTGAATGCGGCGGCTGCGCGGCTCTGAGCGCGCGCGCTTGTATTGCGCGCG 13950

Eco47 III

ATGGTCTCGGTAGCGCTCGCGAGCTTTCGCTTGTGGGCGCGCTTCAAGGCTTGGCTTGCATTGACGAACCGCTCGCAGCAATCGCTGC 14040
 GATTCATATGCGGTGCGGAGCGCGGACAGGTAGAGACGGTGGCTACGAGCGAGCTTGGGCGGAAGCGCGCTCAAGGCGGCTCGATGGAT 14130
 TTGAGGGTGGCAATCAAGCGGCGCTCTTTCGAAGAAGAGACCGCTGAATCCATCTGAGCAACCGCGGCGGCGCATGATGGTGGCAGC 14220
 ATGTTGTGTAATCATCATTCGAGAAATCGCGGTAACCGGAAAGATGCGCGCTTCCCAACAGAGCTGCGTGGAGATCGCGAGCGCC 14310
 TCTCGCGCTTCGGAACGTCACGTTGGCGAGCGCTTCGACAGCCAGCTGTGCGCATGAGCAGCGCTGGACCAAGCAAGAAAGAGCG 14400
 TGAATGCCCAATACCCAGCGGCGCGCGCGCTTCCGAGCGCGCGCTTCCGCGGAGGCGAGCGAGCTGCGGAGTGCATTTCATGAT 14490
 TTTGGCTTTCGCTTGGCGGAGCACTGGTGGAGGATGCTCTCGATGGAGCAGAGATGCGCCATTTCATGGGTTCGAGGAACTACCT 14580
 TTCAGGCTTTCGCGGATCATTCGTAATATGCGAGCATACCGCTTTCGAGTAAGTCTGAGCGCGCTTGGGCGGAGAGCCGGTGCACA 14670
 TGGAAATAGCGCTGCTACTCGCTTCCGCTCCGAGAGGCTCCCGGAGTGTCCGATCTGCGCTTGGCGCTGCCATCGCTCGAGGATC 14760

5 GGGACGGCGAGGGGGCGAGATGCTCGCGAGCACCGCCAGGTGCGCCCTCGGCGCATGCGCCAAAGACCTCGAAGGCCCTCGGCTGG 11850
 ACCTTGGAGGTCTCGCCCATTCGACGGGCTCGATGGCTTCGGGCGACGGTAGATGGATGATATTTAGCCATGATTGCGCGAAGATT 11940
 Bgl II
 10 GCGGCTGCGTCGACAGATCTTTGCGGAGCGGAACGECATTTCCACITGCTCTGGCTCTCAATATTGAATTGAGCCCTGGCGACTGCCATA 15030
 GGGCCAGTCTGTCGACACAGTGTACGGAGCGGCCGATGCTTTCTCCTTTTTAGTCTGACCGAATACTTCTGTTGGCGCCAAAGAT 15120
 CCGTTGCGGAGACTGTCCGGCGAGATGTCGTGTGGAAGCTCCGACGTCAGCGGGCCCATGCGTTGCTAGAGCATAAAACGGTTGCA 15210
 TGCTGGTTCGAGAGGGAGACGCGAGGAGCTCCCTTTGGGACGGATGAGGAATTTCTGTACCGAAATCTCGGAGGAACAGCGGCGAGA 15300
 AGCGGCGCATCGATGGGAACCATGGTTACGAAGACATTGATGATAATGTGACGCAATCGCAATCGTCGCGATGAGCGGCGCTTCCC 15390
 15 CGGCGCGAGAAACGTGAGGAGCTGTGGCABAAGCTCCGCGCTGGCGTGAATGCGTGTGACCTTCACAGAGCGCGAGGCGCTCGCCCG 15480
 GGGGGTGAGCGCGAGATGCTCGCAATCCAGCTACGTGCGCAAGCGCGCGCTCGACGGCGTGGAGCTCTTCGACGCTCGTTCTT 15570
 BamH I
 20 CGGGTTCAGCCCGAGGAGGAGAGCATGGATCCGACGAGCGCATCTTCTGGAGGTCGCTGGGAGGCCCTCGAGCGCGCGGTTA 15660
 CGACCCCGATGCCATTTCGGGCTATCGGCGCTTCGCGCGGAGCGCCCGAGCGGCTACCCTCCCTGCGCGAGTCCGACCGCGAGAT 15750
 CCTAGCGCGCTCGGCCACTACCACTGACGCTGAACAACGACAGGATTATCACCACACAGCGCTCGTACAGCTCAATCTGCGGGG 15840
 25 CCGGAGCGTGTGCTGACAGCGTCTGCTGACCTCGCTCGTGGCGTGGTCAAGCGCTGCGAGGCTGCTCAACCAAGAGTGGGACAT 15930
 BamH I
 30 GGGCTCGCGGTGGCGTGGGATCCATCGCATCAGCGGAGCGGCTATCTGTATCAGGAGAACGCGATCTCTTCGCGCGATGGGATTG 16020
 CGCGGCTTCGATGTGGCGCCAAAGGCGACCGTGGGCGGAGTCCCATAGGCACTGCTGCTCTGAAGCGGCTCGCGGAGCGCTCGCGA 16110
 CGGCGACCACTGACCGGATTCGAGGAGCGGCGATCAACAACGACGCTCGAGCAAGATCGGTTACACCGCGCGAGCGTGCAGGG 16200
 35 GACGCGCGAGTGTGCGCATCGCGGAGCGGCTCGCGCGGCTGGAGCGGATCAGATCAGCTAGCTGAGGCGGAGCGGAGCGGAGCGG 16290
 GCTCGGCGATCCCATCGATCGAGCTCGAGGCTCACCGCGTGTTCGCGCGAAGACCGCACGAAGGAGTTCTGCGCCATCGGCTCGCTCAA 16380
 GACCAACCTCGCGCACCTCGATGCGCGCGCGGCGTCCCTCGCTGATCAAAACGGTCATGGCGCTCGAGCACCGCGAGCTGCGCGGAG 16470
 40 CCTGCACTTCGAGCGTCCGAATCCGAAGCTCGAGCTGGAGAGCAGGCTTTCTACGTCAACACCGGCTCACTCCGTGGCAGCGGCGAG 16560
 AGGTCCGCGCGCGCTGGCGTCACTGCTTCGGCATCGGCGGACCAACGCGCACGTGGTCTCGAAGAAGCTCCGCGCGCGCTCCGAG 16650
 CGGCGCTCGGCGGCTGGCAGCTCCTCACCTCGCGGCTCGCTCGAGGCGCGGCTCGGCGGCGCACGCGCGCATGATCGAGCACCT 16740
 CGATCGCACTCGGCGACATCGATCGCGGATGTACGTACAGGCGACGTGGGCGCGCGGCGCTGGCGCTTCGCGGAGCGGCTCGTGG 16830
 45 CGAGAGCGCGCGGATCTCGCGCGCGGCTCGCGAGCGGCTCGCGCGCTCGATCTCGTCAAGCGCGCGCGCGCGCGCGCGCGCGCG 16920
 CGTCTTCTGTTCCCGGTCAGGAGCGCGACGCTCTTATCGCGCGGAGCTGTACGAGGTCAGCGCGATCTTCGCGGAGCGGCTCGA 17010
 50 ECGTGGCGCGAGTCTTCGCGCGCGGCTCGCGCTCGATCTCGCGAGGCTCTTACCGCGCGAGGCGCGCGCGCGCGCGCGCGCGAGCA 17100
 CGAGCTCGGTAGGACCGCGATCGCGCGCGGCTGTTCGCGGAGGCTCTCGCTCGCGAAGCTGTGATCGGCTCGGCGGATCGTCCC 17190
 CCAGGCGATGATCGCGCACAGCTCGCGGAGTTCGCGCGGCTTCTCTGCGCGGATCTTCGCGAAGAGCGCGCTCGCGCTCGTGG 17280

CTCCGCCCAACGACCAGGGACCTATGTGCTCCCGAGACCGAGACCGAACGGCGCATCGCCGCGATCTGGCAGGATCTCTCGGCT
 18710
 CGAGCAGGTAGGCGCAACGACGATTTCTCCAGCTGGGCGCCCATTCGCTGTTGGCCACGCGAGTCTGTCTCGGCTCTCGAGACCT
 18800
 5 CAAGGTGGGGATCTCGTTGCCGAGTTCTTCGATGCCCGACGGTCGCAAGGGCTTTCCGCTGGTGGACGCGAGCCGGGCGCAAGGCG
 18890
 CGGACCCGTCGCGCGGCAATCGGCCGTGTGAGCGAGAGCTTACCGAATCAAGCCGCCCGGCGCAACAGGCGCGCGCAACAGGCG
 18980
 GTAAACAAGAGGGATCGAGTCATGGAACCGTCGGCGCGTGGACATGAATCAGCCCGCAAGCAGCAGGAGACCTGCGCTTTCCGAC
 20070
 10 CTCTTCGCGCAGCGCGGCTCTGGTTCTTCGACAGCTCGAGCGGGGAGCGCGCTTACAACATGCCCGCTCTCTCGGACGCGCG
 20160
 GCGGTACGAGCTCGACTCGCTGTGCGCAGCGTGAACGAGATCGTGGCGCGCCACGAGTCGCTGCGCAGCAGCGTGTATCATGATGG
 20250
 CGAACCCTGCGAGGTGATCGCCCTCTCGCTGCGCATCGAGGTGCCGCTCGTGGACCTGAGCGAGATCGACGAGCGGAGCGAGAGGCGGA
 20340
 15 CGCCCGCGGCTCATGCGGAGGAGAGCGCGCGCCCTTCGATCTCACCGCAGGCGCGCTGCTCCGAGCCAAGCTGCTCCGCGCTCGGCGA
 20430
 GGCGGATCAGCTGCTGATCTTGACGATGCATCATATCGTCTCCGACGGCTGGTGGATGGACGTGCTGTTCAAGGAGCTTTCCACGCTTA
 20520
 CGCCGCTTTCACGAGGCGCGCGCTCGCGCTCCCGGAGCTGCGGATTAATACGCCGACTTCGCGGTGTGGCAGCGGAGCTGCTCA
 20610
 20 GCGCGAAGTTCTGGAATCGCACCTCGGCTACTGGAGAGACCACTCCGCGCGCGCCCGCGCTGCTGGAGCTTCCGATGGACGGGCGCG
 20700
 GCGCGCGCGCAGAGCTTCCGGGCTCCGAGCGCGCTTCGACTCCACTCTCCCTGCAACAGCGCGTCCAGCGCTCAGCGCGCAGGA
 20790
 AGGCGCGACCCCTTCATGACGCTGCTGACGGCGTTGAGCGCTGCTCTCGCGTTATGCGCGCGCAGAGCGATCTGGTGGTGGCACCC
 20880
 25 CATCGCGAATCGACCCGAGCAGAGCTGGAGGGGCTGATCGGCTTCTCTGTCACATGCTGGCGCTGCGCATCGACTCGGGGCGGACCC
 20970
 GAGCTTCCGCGAGCTGCTCGGGCGGCTGCGGGAGGTGACBTGGGCGCTTACGCGCACCAAGGACCTGCCCTTCGACGCGTGGTGGAGGA
 21060
 GCTGTACAGGCGGAGCGCCAGCCACAGCCCTTGTTCAGGTGCTCTTACGTTGCAACACCCCGATGGATGGCAGCAACAGAGC
 21150
 30 AGACATTGCATCGGTGGCGCGCGCTGGTGAATGAAGGCGCGCAATTCGATCTGATCTGGAGCTCTCGCAATCGCCGCAAGGGTT
 21240
 GCTCGGCACGTTGAGTACAACACCGACCTGTTGCAAGCGCGCACCATCGAGCGGATGGCGCGCCACCTGGAGGTGCTGCTCTCCAGCGC
 21330
 CGTGGGCGCGCGGATCGACCATTTGCGGAGCTGCGCTCATGGGGCGCGAGGAGCGCAGTGGGTATTGGTGGAGTGGACTCCACTGC
 21420
 35 CGCCCTGTATCCGAGGACCATTCGATGCAAGGCTGTTGCAAGCAAGTGGAGCGGTGCGCGAGGCGACCGGGTCTCTCCAGCA
 21510
 GCAGAGCTTCACTATCGAGAGCTGAACATGCGCGCAATCAGCTCGCGCATCAGCTGCGGAGCTGCGCGTGGCGCGTGGCGCGAGGTGCGCGT
 21600
 40 CGGGTTCTATCTCAACGGTCAATCGAGACGGTGGTGGGATCTTCGGCGTGTCAAGGCTGGCGGGGCTACGTGCGCTCGATCCGAC
 21690
 GTACCCAGCGAGCGCTCGGGCTCATGATGGCGGACGAGCGCCCTCGGTGCTGCTCAGCGAGCGCTGCTCTCTGAGGTGCGCGC
 21780
 Eco47 III
 45 CCACGCGGATGCAACGCTGGTACAGCTCGACGCGCTGCAAGAGGCTCTCCAGGCTGCCACACCATACCCCGGAGCGCGCTCACCGC
 21870
 CCAGAACCTCGCATACGTCATACACTTCGGGCTCGACCGGCGGCGCAAGGCGGTGCTGCTGAGCACCGCGGCTCTGCAACCTGCC
 21960
 CACCGTGCAGGCGAAGCTCTATGGAATCGCGCGGCGACAGGCTCTCTCAGTTCCGCGCGCTCTGCTTCGACACATGCTTCTGCGAGAT
 22050
 50 CGCGCTCGCGTTGCTCTCGGGAGCGACGCTGGTTCATGGGACGCGCGGACGAGCTTCTCCGCGACCTCGGCTGGTGGAGCTGCTGAAGAA
 22140
 Eco47 III
 GCACGCGGTACGCGGATGCTCTGGCCCTACCGTGTCTCGACGCGTGGCAGAACACAGAGCGCGCGCTGCGGCTGCGCGTGTCTAC
 22230

DE 198 46 493 A 1

<p>Eco47 III</p> <p>GATGSCGGTGAGGCGTGCCCGGGGAGCTCGTCAAGCGCTGGAAGGACCCGACGGCCCTGTTCAACTCCTATGCCCCGACCGAGAC</p>	22320	5
<p>BamHI</p> <p>GACCATTTGGCAAGCTCCGACGGACCTGTCGACGAACGGATCCCGCCCATCGGCCGTCCGATTGCCAATACGCAATCTACGTCT</p>	22410	
<p>Eco47 III</p> <p>CGACGAAGCGCTCGAGCCGGTGCCCATCGCGTGCCGGGCGAGATCTTCATCGCGCGGTGGGCGTCGCCCGGAGATACGGGCGTCC</p>	22500	10
<p>Bgl II</p> <p>GGACCTGACGGCCGAGCGATTTCGTACCGACCCCTTCGGGCAJACCAAGGGGCGCGCTGTATCGGACGGCGATCGGCGCGCTGCT</p>	22590	
<p>GCCGGACGGAACCTCGAGTTTCCTGGTCGAAACGACGAGCGAGGTGAAGGTCCGCGGTCTCCGCATCGAGCTGGAGGAGATCCGCGCGG</p>	22680	
<p>GTTCCTCAAGCACC CGCGGTGCTCAAGCCGTGGCGTGGTGC CGGAGGACACGCCGGGCGACAAGCGGCTCGTCGCGTATGTCCTCG</p>	22770	15
<p>ACCGCGAGGAGCGCGCTGACCGCGCGGAGCTCGCGCAGTCCGTGAGCGAGCGATTGCTTGCACCAATCGTCCCATCGTCTTCGTGGC</p>	22880	
<p>GCTCGACGCTTGGCCCTGACGCGGAATGCCAAGGTGGACCGCGCGCTGCCGGAGCCCGAGCAGAGCGCGCGCGGCGAGGACCACT</p>	22950	
<p>CGCGCGCGCAACCGCTCGAGGAGGAGCTCGCCAGGATCTGCGCGAGCTGCTCCGGTCGAAAGCGTCCGCGCTCCAGCAACTTCTT</p>	23040	20
<p>CGAGATCGCGCGCACTCGATCTGAGCATCCAGATCGTGGTGGCGCGCAGCAGGAGGCTCGCCCTCAGCCCGCGTCAGATGTTCCA</p>	23130	
<p>GCACCAAGCATCGCGAGCTTTCGACCGTGGCTAGAGCGCTGAGGCGGTCCACGTCGAGCAGGACCCGGTGACCGGTCCCGCGCGCT</p>	23220	
<p>Eco47 III</p> <p>CACGCCGGTGACCGCTGCTGGCTGAGCAGGAGGCGGCCGAGCGCGACCACTTCAACCACTGATCTTCTCGAGGTACCGAGCGGCT</p>	23310	25
<p>CGACGAGAGCGCGCTGAGCAGGCCATCGCGCATCTGATCGACCAACGACGCGCTCCGGTTCGCGCTCGCGCGCGAGCAACCGCGCG</p>	23400	
<p>Bgl II</p> <p>CCACCAGGTCTTCGCCGCGCGGAGGCTCGACCCCAATTCAGCGCGTCGACCTCGGGCGCTGCCAGCGCGAGCAGATCTCCGCCAT</p>	23490	30
<p>GGAGAAGGCCCGCAGCGAGGCGAGGCGAGCTCGATCTGGCGCGCGGCGGCTCGTCCGCGCGTGTCTCTCGACCTCGCGAGGTCGC</p>	23580	
<p>BamHI</p> <p>CCCGCAACGGCTGCTCGTCATCGCCCAACCATATTGCGGTGACAGCGTCTCTTGGCGGATCTGCTCGACGATCTCTTTCGGGGCTATGA</p>	23670	35
<p>Eco47 III</p> <p>GCAGGCGCGCCCGCGGAGGCGTACCGCTGCGCGCCAAGACCACTGCTGTAAGCGCTGGGCGGAGCTGCTACCGAGCACGCCGCTC</p>	23760	
<p>CGAGGCGCTCAAGCGGAGCTCGGCTACTGGCTGACTCATCGCGACGAACGGTAGCTCCGCTGCCCGTGGATCGACGGGCGGCGAGGA</p>	23850	40
<p>CGTGTGGGCTCGCGCGCCACATCGTCTGCTGCTCACGCGGAGCAGACGGAGCAGCTCTGCGCGAGGTCCCGCAGGCTACCGAC</p>	23940	
<p>ACGGATCGACGACGCGCTCTCACTGCGTTCGCGCAGGCCATCGCTCGGTGGACGGCTCGCCCGCGGTCTCTCGACCTCGAGCGGTCA</p>	24030	45
<p>CGGGCGGAGGAGCTCGCCGGGTAGACCTCACGCGCACGGTGGCTGGTTTACGCGCATGTACCGGATCTTCTCGCGCTCGACCGCGC</p>	24120	
<p>BamHI</p> <p>GGATCCGGGTGAGGCGCTCAAAATCGATCAAGGAGCAGCTCGCGCGCTGCCAGGCGCGGGCTCGGCTACGGCTTGTTCGCTTACCTTCG</p>	24210	50
<p>GTCCGATACCATCGCGAGGTTCGCGCTTGGCGCAGGCGAGCTCTGCTTCAACTAECTCGGCCAGCTCGATCAGGCGATCCCGAGCC</p>	24300	
<p>TGCACCGTTCGCGCGGCGCGAGTATCAAGGCTCGGAGCGCAGCCCGCGGCCCATCGCGGCCACCTCATCGAGGTGAACGCGAGCAT</p>	24390	55
		60
		65

	CGCCAATGGGCGCTGTACGCCACGTGACGTACAGCGAGCGCCGACGAGCCGAAACCATCGAGCGGTCGCGGCGAGCTTCGTAC	24480
	CGCGCTCCGCGCCTCATCGCGCACTGCACCTTGCCCGAGBTGCGCGCAACACGCTTCCGACTTCGACAAGGTGCGCTCGCGCAGGA	24570
5	GACCATCGATGCTCTCGACGCAATCGACGGGGCCCCGGGCCGTCTGCGAGGGGAGCCGATCGAAGACBTCTACCCGCTCTGCGCGT	24660
	CCAGGAGGGCATCTGTTCACACGCTTACGCCACCGATTACACGGCGTATGTCCAGGAGTTCCTACTGGACCGTGGAGGGCGATTTCGA	24750
	CGCCGAGGGCTTACCCCGCGCCTCCAGGACGTGCTGCTCGGCATGCCGCCCTGCGCACGTGCTTCGCTGGGAGCGCCTCGATGCTCC	24840
10	ACTTCAGATCGTCCGACCGGCGCGGTCTCCCGCTCGAGCACCAGGACCTACCGGCCCTCGCGCGGAGGACGACCGCGCACATCTC	24930
	CCGTACGTCGAGGCGAGCGCCAGCGCGGTTCGATCTCGAAAGGGGCCCTCATGCGCGCGGGCTGCTCCGGCTCGCAAGGACGC	25020
	CTGTGCTCTGTCGAGACCATCCACCACCTGATCTTGGACGGTGGTGCACAAATCTTGCTCAAAGAAGTTCACGCTCTACGAGC	25110
15	GCACCGCGGACACCGTGGCATCTCCTGCTGGAGCTCGAGCAGCGCGGCCCTACGGCGATTACATCGGCTGGCTCGCGAAGCAGGACCA	25200
	GGTGGCACCGCGCGCTTCTGCGCGCGGAGCTCGAGGGCTTCTCGCGCGGACCGCGCTCGGCGTGCACCGCGCTGCTGCGCACGACGA	25290
	CGGCGCGCGCGGTTGGTTCGCGCGCATCGCCCTCTCGGGGACGACGCGCGCGCTCGCGCGCTTCGCGCGTCAGCATCAGCTCAC	25380
20	GATGAGCAGCTGTTGCAAGCGCGTGGCGCTGCTCTTGTCACGCTACAGCGCGATCCGACGCTCTTTCGGTATGACCGTCTCGGG	25470
	CCGCTCGCGCGGATTCCCGGTATCGAGCGCATGACCGGCTCTTCATCAACACCATTCGCGTGGCGTGGCGGAGCTCGCGACGCGTC	25560
	GGTGTCTCGCTGCTCAAGCGCTCCAGGAGCAGGAGGAGCTGCTCGAGCAGGACAGCGCGCTGCTGCGAGTCCAGGCGCATAG	25650
25	CGACGTGCCGCGGGACCCCGCTCTTCGAGACCTCGTCTGTTGAGAACTACCGGTGCGAGTCTCTTCAGCGCGCTTCGCGTGA	25740
	GGGGCGGACGCGCGGAGGAGGCGCTCCGATGATCGATGCGCGATATATGATGATCCACCGTATCCGTCGAGGTCGTCCGCGCTT	25830
	CCATGGGACGCTTATCTCAATATTGGCTACGAGCGCGCGGTTTCGACAGCAGCGCTCGAAGGATGATCGGGCAGTCAAGCAGCT	25920
30	GCTCCGGGCTTCTGTCAGAGGCGCGGACGTCGGTCCGCGATCTCGCGTGTCTGACGGCGAGGAGGAGCGACCGAGTCCACCGCTG	26010
	GATGCCACGGCGCGCGCTATCCGAGGCGCATTCATGACGAGCTGTTTCGAGCAGCAAGTGGAGCGGTCGCGCGAGGCGAGCGCGT	26100
	GCTCTCCAGCAGCAGCTTGAAGTATCGAGAGCTGAACATACCGCGCAATGAGTGGCGCTACCTTCGCGAGCTTCGCGCTCGGGCC	26190
35	AGAAGTCCGCTGGGCTTGTGCTCGAAGGTCGATCGAGACGGTCTGCGGATCTTCGGCGTCTCAAGGACGGCGGGTCTAGTGCC	26280
	GCTCGACCCGACGTACCCGAGCGAGCGCTCGGGCTCATGATGAGGAGCGCGCGCTTCGGTCTGCTCAGCGAGAGCTCGCTCTCTC	26370
	Eco47 III	
40	GAAGCTGCCGCCCCACGGGATGCAACGCTGGTACAGCTGACGCGTGCACGAGCGCTCTCCAGGCTGCCACACCATACCCGCGGAG	26460
	CGGCGTCACGGCCAGAACTCGCATACGTATGTACACTTCGGCTGACCGCGCGGCCAAGGGCGTGTCTGTCGAGCAGCGCGGCT	26550
	GTGCAATCTGCCACCGTCAAGGCAAGCTCTATGCAATCGCGCGAGCGACCGGCTCTCCAGTTCGCGCGCTCTGCTTCGACACATC	26640
45	GTCTGCGAGATCGCGCTCGCTGCTCTCGGGAGCGAGCTGGTGTGAGGACGGCGGACGAGCTCTCCGGGACCTCCGCTGGTCTGA	26730
	Eco47 III	
	GCTGCTGAAAAGCAGCGGTCAGCGGATGCTCTTGGCCCTTCCGTGCTCGCAGCGCTGCCAGAACACAGAGCGCGGCTTCCCGCT	26820
50	Eco47 III	
	GGCGGTCTGCGGATGGCGGCGAGCGCTGCCCGGCGAGCTCGTCAAGCGCTGGAAGGACCCGCGCGGCTGTTCACCTCTATGG	26910

CCCGACCGAGACCAACCATTTGGGCAAGCTCCGACGGACCTGTCCGACCAACGGATCCGCCCCATCGGCCGTCCGATTGCCAATACSCA	27000	
AATCTACGTGCTCGACGAAGCGCTCGAGCCGGTCCCATCGCCGTGCCGGCBAGATCTTCATCGGCGGCGTGCGGCTCCCGGGGATA	27090	5
TCACGGCGCGCGGACCTGACGGCCGACGATTCGTACCGACCCCTTCGGGCAAAACAAAGGGCGCGCCTGTATCGSACCGCGATCG	27180	
GGCGCGCTGGCTGCCGACGGCAACCTCGAGTTTCTCGGTCCAAACGACGAGCAGGTGAAGGTCCGCGTATCCGCATCGAAGCTGGAGGA	27270	10
GATCCGCGCGGCTGTCTGAAGCACCGCGGTGCTCAAGCGTGCGCGTGGTGGCGGAGACGCGCCGGGACAAAGCGCTCGTCC	27360	
CTATGTCTCGGACCGGAGGAGCGCGCTGACCGCGGGAGCTGCCCGAGTCCGTGAGCGACCGATTGCCCGGACCATGCTCCCGTC	27450	
GTCCTTCCTGGCGCTGACCGCCCTGCCCCACGCCGAACGGCAAGGTGGACCGCGCGCGTCCCGGAGCCCGAGCGGAGCGCGCG	27540	15
CGAGGACACGTGCGACCGCGCAACGCCATCGAGGAGGTACACCAATCTGGGCGACGTACTTGGGCAAAAGCGGTCGGTGTGCA	27630	
CGACAATTTCTTCGATCTCGGCGGCTATTCCTGCTGCTCGTCCGGGTGATGATCGGCTCGGCGAGCGGTTGATCGCCCGCCCTCGAT	27720	
GGTCGACCTCTTACCTATCGGCGGTGGCGTGGCTCGCGCGGTTCCTTGGCGAACGGCGCAAGCGCAATCGCCGAGGAGCGCG	27810	20
GGCGGACCTACGGAGCGCGCGCGCGCTGGAGCGCGCGCGCGCGCGGCGGCGGAGGCCATCCGTGGCCCGACCTGACCGGCGACCC	27900	
TTCAAAGCCCCGCGCTTCTCGCACATCGCGCGCTCGAGCGCGCGCTCCAGCGCGCGCTTCGCGACGAGGAGCGCGGAGACGCGT	27990	
CCAAGCGCTTCGTGGGCTCTTTGCCCGCAATCGGAAGCTGGCGCGCGCTTCGTGCGCCCTGCAATGCTGCCATGTAGAGCTCTCTCG	28080	25
CTCGCCGCGGCTCTTTCTTGGCGGCGCTCGCGGATGACCTCACATCTGATCCCTTCTTGGCTCGTCCAGAAAGTATTGACGSCCA	28170	
GGCGCGCGCTTGAGATCTTCCGGCGCGCGCGATTTATCTGCTCGCGCGCGCGTGAATCTCACCTGCGAAGGATTATAATGAACATA	28260	30
ACATTGGGTGGCTTCTACCGCGCGCTCGCGAGCGCTTGGCTTGGTCCCGCGCTGCGAGCCGAATCACCGTGAGGATGCGCGCTTCTGTA	28350	
CGTCAGCAGAGAGCGCGCGCGCGCGCGCTGAGTGGTGGCGCTCGCGCGCAAGCTCCAGCGCGCGCTGAGCGCGCGCGCGCGCGCG	28440	35
AAAAGGCTCCGGCAGCCGCGCGCGCTCGAGAGCGGGACTGTGTGCGCGGCGCGCCACGGGCGTCTCGGACCTGGTCCGAGCAGCG	28530	
CGACGAAGCTTGAGATCTCTTTCGGATCGGACGATCAACGAAGACCTTCGTCTCTACGCTGATACTCATGCTCCGGGCAAGAGCGCGGT	28620	40
TGTGCTCGACGACCGCGTGTGGAAGTATGTGAAGGCGATCCCGCGCGCGGACGAGATGACGCTCGCGCAGATCTTGGTCAACGAGCG	28710	
GGCTCTTCGATTACAGTACAGCCCGCGCTCGGCGCAATGATCGAGGTGGATCCGACCCCGCGCTTCGCGCGGCAAGGCTCATCGCCC	28800	45
TCGCCACGGCGGAGCGCGCTATTTTCGCGCGCGCGCGGTTTTCGCTATTGGAACCAATTACATCGTGGCGCGCTTGGTGGCGAGG	28890	
CGGTGTTCGGCGGAGCGCTCGCGCGCTGCTCGGACGCGATCTTACACCTGTGGGCTCGCGCACAGCTATCTGGAGCGCGCGAGC	28980	
CGCGGCTCAAGGCTCATCGCGCTACGGCGACTACGGCGCGGCTTGGTTCGACATACCGACGAGCTGTGCGCCACCGAAGCGTGGG	29070	50
CGCGCGCGCGCTTGGTTCGAAGCTCGATGACCTCAATCGCTTCTTTCGCTGCTATCAGCCAGCGAGCTGCTCTGCTCGAGAGCTTC	29160	
AGGACATGACCACTGGACCCGACGATGTGGCCCCACGAGCCCGGATATGCCCTCGGCTCATCGAGCGCGATTCTGCGCTCGGCTCCG	29250	55
		60
		65

TCAACGGGCACTGCGGAATCATCTGGGGCTTTCAATCGGGCTCGTACGGGTGCCCCGGCGGGGACGCGATCACCGGCTCATCAACC 29340
 Eco47 III Eco47 III
 5 GGAGCGACGCCGACGCGAGCGGGCTCTGACGAGCTCCCAAGGTCTGTAAGAGCGCTGATCGAGCGGAATCGGAGCGCTTCGGGCGG 29430
 GTGGTGATGGCGCCCGGCGCTCAGAACCGGACGCGAGCCCCGCTCAGCGGGCTGCGCCGGGGGACGCGGCCACGCGGCCCGGACCG 29520
 ACGAGGAGCGCGGACGCGGGGCGCTCGGCGGCTGCTCGCCGACCCGCGCTTGCCGAACACGTAGAGCGGCGAGGCGGACGCGG 29610
 10 ACCCGGGCACCCGCGAGCGCGTGGCGATCGGCACTCGGACGCTCGGCGCGCGCGGCTGCTGTGCTGCTGGCTCGGGAAGACC 29700
 AGCACCGCGCGCTCAGGATGGCGCGCGCCAGGCTCGTGAGGACGAGCGCCGAGATCACCATGACCGGGCTGTTCACCTCGCTGCT 29790
 CGCTCGTCGAAGTCGCGGAACGCGGCCCTCGCGCGCGAGCTCCAGCTCGATCGCGGCTGCTCGGCGCGCGCTCGTCCGCGAGCGG 29880
 15 ATCTCGTBGACGCGCGGGCTTGGCCCTCCAGCGCGCGGATGCGCGCTCTGTCGCGGCGCGGCTCTCTCCACGTCGCGCCCTGGCGGG 29970
 ACCCGCGCACGCGCGCGGACAGAGCGCGCGAGCGCGGGTCGAGCGGGCGCGCGCGGGCTCGCGGGCACGGAAGCGCGCGG 30060
 GCGCGGGAGGCGCGGGCGGCTCGCGGGCACGGAAGCGCGCGCGCGGCGGAAAGGCGCGGGCGACTCGCGGGCACGGAAGCGCGGG 30150
 20 GCGCGGGAGGCGCGGGCGGCTCGCTGCAACGAGAGCGCGGGCGCGCGGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30240
 GCGCAACCGCGCGCGGGCGCGGAGCGGAGGCTGCTCCATGCGCGCAGGCTCGCGCTCGCACAGGGCGGGTCAAGCGCGCGG 30330
 AGTCGAGAGCACGAGACCTCGCGCGCGAGGAACAGCG 30420
 25 CGCGCGCGCGCTCCCTCCCGCGCGCGCGCGCTCG 30510
 TGCGCTCCATGAACGAGCTGAACGCGAAGTACGGCAGCAGCTGCGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30600
 CGCGCGCGCGCGCGCGCTCG 30690
 30 CGCGAAGGACCGCGCGCGAGGAGCTGAGCTCGAGCAGCG 30780
 GCGCGCGCGCGCGCGCGCTGGGCGTGAAGTAGGCGTGGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30870
 TCGCGCGCTGCTCGCGCATGCTCCAGAGCGCGCGCGCTCG 30960
 35 CGAAGATGGCGAGCGCTCTCTGCGTGGCGTGGTGGCGGGCG 31050
 CGGTGAGCGCGTTCGCTCGACCTCGTGGTGGCG 31140
 TCATGCGCGCGCGCGCGCTTGGCGGTGACG 31230
 40 GCATGTCGCGAGCGCGCGCGCGTCCGCGCGCTCGCTCGGGTTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 31320
 GCTCGCGCGCGCGCGCTTCTGAGCGAGCG 31410
 GGGTGGCGCGCGCGCTCGAGCGCTCGCGCGCGTGGATCTGCG 31500
 45 TGGCGCGCGCGCGCTCGCGCGAGCTCG 31590
 Eco47 III
 CGGCGCGCAAGAGCGCTCTCCACCTCG 31680
 50 CCGGATGAGCTGCTCCAGCAGCGCGCGCGTGGTCTGCTGAGCG 31770
 ACGAGGCGCGGATTGCGATGTTCCGTGCTTTTTGAGCGCGCGCTGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 31860

AGCGAGGGTGTTCGCGGGCTGCTCGAGAGCTCCCTTTGCCCCACACCCGAGACACGAATGCTCCGCGCCGAGCGGGTTGACCGTGCAC 31950
 CCGCCGGAGAGCCGATGATACGGTCCGGCCGATGTCGGAGAGTGTAGCTCAACTCGAAGAACACCCGCGCGGCTCACCAGGACACTGCTA 32040
 CCGGATGCTGGGTTCCGTGCTCGACGCCGACACCCGCTCCAGGAGACGATGGTCCGCGCCCTGGCGGAGCCTGGATAAGTTGACGGGGC 32130
 CTCGTGCTGCGCACCTGGCTGTACCGCATCGGACBAACGTCTGCATCGACCTGCGGGCCGACCCGCGCGCGCCGGCGCGCCCCCATCGA 32220
 GGAAGGCCCGGTGCGACCGTGGACGACGCGCTCGAGACGCGCCCGCGCACCCACTGGCTGGAGCCCGTCCCCGACGCGACGCGCTGCG 32310
 GCGCGACATCGACCCCGCGAGCGGGCGATGCTCCGCCAGAGCATCCGCTCCGCTTCGTGCGGGCGCTCCAGCACCTGCGCGCGAAGCA 32400
 GCGCGCCGCGCTGCTGCTACGGAGGTGCTCGGCTGGTCCCGCGCGAGGTGCGCGACAGCTCAACACCTCGGTGCGCGCGATCAACAG 32490
 CGCGCTCCAGCGCGCGGGGCGACGCTCGGAGCGCGGATCTCGCGGACGCGCGCCCTCGCTGCGCGAGCGCGAGTCCGCGCTGCTGCA 32580
 CCGCTACGTCAACCGCTTCGAGCGGTACGACGTGACGCGGTACCGCGCTGCTGACCGAGGACCGACCTGTCGATGCGCGCGCTTAC 32670
 CCTGTGGCTCCCGCGCCACGAGTCGATCCGCGCTGGCTGCTGGGCGCGGAGCGGGTCCGCGGGTCCGCGCTCATCCGACGCGCG 32760
 GAGCGGCTCGCCCGCTTCGCGCACTATCGCCCGCGCGCGAGGGCGGCGACCGGGCTCGCGCGCTCATCGCTCGAGCTCGCGGGGA 32850
 CCGCATCGTACGATGACGCTCTTCTCGACACCGAGACGCTCTTCCCGGCTTCGGCTTCGCGCTCGATCTACCGCGTAGCCCGCGCG 32940
 GCGCTGCTGCTCGCGCGGGTGGCT 33030
 CCAGCGACGGGGCGACGATTTTTTTCTGAGCGACGATGAGTCTGACCGGGCGGGGGTCTACGGGGTGAATCCAACCGGAGGACAC 33120
 CATGACCGTGACCATCGCCAGCATCGATCATGTCGACGAGGACCTCATGACCGGGCGCCAGCGCAAGGGCGCGCGCGCGCGCGCGCG 33210
 CGACGCGCGCGCTCCAGCGGAGCGGTGTGGCGGGCGCGCTCTGAGCGGGTGGCCACGCTGTTCTGACGTTGACGCGCGCGGTCAA 33300
 GGTGCTGAAGCTGTTCGCCCGGAGGCGTCCAGCGCGAGCTCGGGTTCGGCGCGACCTTCGTCGCCACCTTCGGCTACCTCCAGATCG 33390
 TTGCTCTGTCGCTACCTGATCCGCGCACCGGGTGTGCGCGGATCTGTGAGCGGGTACCTGGCGCGCGGATCGGATCCAGCT 33480
 GCGGGTCGAGAACCCTCTTCAGCCACACGCTCTTCCCATCTACGTCGCGCGCTTCTCTGGCGGGGCTCTGGCTGCGCGACCGCG 33570
 CGTGCGCGGCTGACCGCGAGCGCGCTGCTCGCGGGCGGATGAGCTTCAGCTTCAAGAGGTCATCATCGCTTAAAGGAGAGCGAGCC 33660
 ATGACCACAAAGAACCCCGCAAGCTCTTCGTAACTGTCCGTCCGCGACCTGAAGCGATCGATGGAGTCTTCAGCAAGCTCGGGTTC 33750
 GAGTCAACCCGAGTTCAGGACGAGAAAGCGCGCTGCACTGCTGTCAGCGAGGAGGCTATGTCATGCTCCCTCGTGGAGTCTCTTC 33840
 Bgl II
 AAGACGTTATGAAGAAGGAGATCTGACGACGAGCAACGACCGAAGGCTCTTCGCGCTCTGTCGACGAGCGGCGGAGGCTCGAC 33930
 BamHI
 GACATGGTGAAGAAGGCGGTGCGCGCGCGGGTCCGACGCGATGATCCGAGGATCACCGCTTCATGTACGGGTGAGCTTCTACGAC 34020
 BamHI
 GTGGATGGCCACCACTGGGAGGTGATGTGGATGGATCCCAAGGCGATCCAGCGTACCGGACGCGGGCTGGGCGCGCGCGCTGGAAGAGCC 34110
 CCGCTGAGCGGGGAGGCGGGAGGATCACCGTCTTCGTAGCCACAGCGATGACGATCCGTCGCGCTTCGATCGAAGCAAGCTGTTA 34200
 CGGGCGCGTCAGAGCGCGTCCGAGGTGATCCGAGCGGACGAGGACGCGGACGAGCGTGGCTCCGATGGAGTGAAGCGAGTCTCG 34290
 CCGATGGTCTCGGGGTGATGAATGGATGAGTAGGGGCTCGCTCTTCGTACGCTGCTTCGACGCGGACGCGGAGCGCGAGCTCGAAG 34380

CCCTCGCCGCTCCGCGTCTGTCGAGCAGCGCTACCGGACAGCGCTGATCGAGGCCATCGGGTTCCGGAGCTCGTGGCGGAGCATCGCGAGAAC	38900
TCGTCTTTCGCGTGTATCGCGGAGCTTCAGCGCCTCGACGAGCGCTCCACCGCGCTCCGGGCGCGCACTGGTCGTCACGTCCAACCGG	38990
AACACGAAGACGCCCTCGACCGCCCCCTCGGATCGCGCATCGGCTGGTAGACGAAGTTGAAGAACCTCTCTCCGTCGTGCCGTCCGCC	37080
CGCGCATCGAGCCGACCGGAGCTCTTTCGCGCATGGGCTCGCCGGTGGCGACCAACCGCTCGAGGAGCTCCAGATGCTTGTCTCC	37170
TCGAGCTCGCGGAGGGCGGCCCGGATGGGCTCGCCACGAGCGATCGACCGCGGACGAGCGCTGGTAGACCGGGTTGACCACTCGAAG	37260
ACGTGCTCGGCCCCGCGGAGGATGGCGATGGGCCCCGGGCTCGCATGAAGAGGTCTTTCAGGTACTGGCGCTGCCCTTCGGCTTCGCGC	37350
CGGCGGCGCGGAGCTCGACGTGGATCGGACCGCGCGGAGAGCTCTTTCGGGAGAACGGCTTCACGAGGAAGTCGTGCGGCGCGG	37440
TCGAGGCTGTGACGCGCGCTCTCTCGCCGCGCGCGGAGAGCATACCAACGGCGACCGCGGGTTCGATGTCGGCGCGCGACGCGC	37530
CTGAGCAGCGCGAAGCGCTCGAGCGCGGATCATCACGTGGTGAGCAGGATCCGGCGGGTGGGCGCGCGCGCTCCAGGGCGGCG	37620
CGACGTCGGCCACGGCTTCACCGTCCACCTCTCCCGACGACGCGCAGCGCTACTCGCGCATGTCTCGCGCTTGTCTCGGCGGACG	37710
AGGACGCGCCCCGCGACGCTTCGCGGCGGCCCCCTCGCCGCGGCGGAGCGCGCGGCTGCTCGCGCGGAGGCACTGCGCGGCTCG	37800
TCGAGGAAGGGCGCGGCTCCCGCCCCCGCGGCGGCGGCGGCGGCGAGCG	37856

5

10

15

20

25

30

35

40

45

50

55

6

6

DE 198 46 493 A 1

SEQUENCE LISTING

<110> Gesellschaft für Biotechnologische Forschung GmbH

5 <120> DNA-Sequenzen für die enzymatische Synthese von
Polyketid- oder Heteropolyketidverbindungen

10 <130> DNA-Seq zur enzym Syn von Polyketid .

<140> DE 198 46 493.2

<141> 1998-10-09

15 <160> 1

<170> PatentIn Ver. 2.1

20 <210> 1

<211> 37856

<212> DNA

25 <213> Sorangium cellulosum

<400> 1

30 ggatcgcggc gccctcggc tgcctctga gcgtgaggag gaactccac gccaggcgcg 60
acttgccgag gccaggcggc cccaccacca ccaccgcgtt cgggaggggc tgcgcgacgc 120
aatggcgcca ctgcgtcgcg agctgcgaga gctcgcgctc ccgccccacg cagggcgtcg 180
gcttgccgag gagccgtggg acggcatcgc gctcctcctt egggcgcgca agccagcacc 240
35 ctccggggccc ctgtaccgctc tcgaagcggc tcgcgagcag gctggcgctc gcgtcgctga 300
gccggatctc cggcgggcgac aggccatctc gcccgcgcat gagctgcgcg acccgatcga 360
ccagctcgcc gaccggcgag ctgcctcga cctcgggcag cctgtcgcg acggacacgg 420
gcacgcctcc gagcgccgc cgcagcgcg gggcgcgagt ggccgcccgt gtggcgagat 480
40 cegtggggcga ctgcggcgcg gacagcgcg cgagccacca gcgcgcttgc agccgatcga 540
ggcgcccgcc gtggcgcgcc gcgatgtccc gcagcgctc ggcccgcgcg gcgcgctgt 600
cctccgagag cgtggcgccg gccctggcgc cgcctctctt ggccaggatg acgcacatca 660
ccttgcgctc ggccgtcggt atcgctcgc ccggcgcggc cggcgccgag acccgctcgt 720
45 ccccgatcga gagccctcg ccggccacgg cggcgagctc cgcgcggcg gggcgccgt 780
cgcgcggcg ctctcccgcg ttcttcgcca gcctccgcgc caccaggcgc tcgagcggt 840
ccgggatacc gtcgcgagc tccccgagcc gcggcggtc ttccaggacg acccgcatca 900
50 ggagcgcgag cgcgctgttg ccgaggaacg gcggcgcccc cgcgaggcac tggaaacagca 960
cgcacccgag cgcgaacacg tcggccccgg cgtcgaccgg cgcgtcgccg cgcacctgct 1020
cggcgctat gtaccggggc gtgcccagca cggccccggg cgacgtgagg gtcggcgcg 1080
gccggaggtg gcgcgcatg ccgaagtga gcagcgtgac gcgctcgacc gcgcgcccc 1140
55 cgagcatcag gttgctcggc ttgaggtcgc ggtgaacgac gccgagccag tggatcgcg 1200
cgagcgtcgt ggccacgcgc gcggccagcg ccacgctctc ggccagcgtg agcgcgcccc 1260
cggcgagccg ctctccagg gtcacgcgt cgagccactc catggccagg tacggccgccc 1320
ctgcgcgggt caccctgct gccacgtact gcaccagcc gggcagccgg agcgtcacga 1380
60 gcgcctccgc ctcccgcgcg aaccggcgca ggtcggtggc gctcgcgccc tgcaagacct 1440
tgagcgcgac cgcctgcccc gacaccgggt cgcgcgcccc gtacacgtcc cccatcccc 1500
cggagacggc gagccgctcg atctcgaaac gatcctcgat cacatccgct gcgcgcatgg 1560

65

cgggtgccaat gtactccgcg cgagcctcgg gcccccgcg gtaagtgcgg ccctgcgccc 1620
 ggttgaacgc cagcccgagc gtgaccgect cgcgctcggg atccacggcc gccggatcgg 1680
 tccacgcctc gacgagcgcc tgcgttgaac aaccgcgcac cgggcgcacg cagccggcat 1740
 cgcgcgctg gccacccggc gctgcgcgcc ttaggetcac ctcgcgatg ccccgctggt 1800
 tcaacacggc aggtccctgc aaccggccg atcaactacat gctcccgccc gaggagcget 1860
 tgcccgagcgt gcgcgatctg gtgcgatcga aggcctactt cgtcctgcac gcccccgccc 1920
 agatcggaac gacgacctcg ctgcgcacgc tcgcccagga tctcacggcc gaaggcgct 1980
 acgtggccgt cctcgtctcg gcggaggtcg gcgcgccctt ctctgacgat cccggcgcg 2040
 ccgagctcgc gatgctcgca gaatggcgcg gcacgcggcg cgcgcagctc cccgcgatc 2100
 tgccggccgc accgttcccc gatgcgcgg ccggtcagcg catcggggccc gccctgcgcg 2160
 cctgggctca ggcgcgcgg cgcgcgctcg tgccttctc cgacgaggcc gacgcctgc 2220
 gcgacgcgac gctcgtctcc ctattgcgc agatccgcag cggctatccc gaccgcccgc 2280
 gtgacttccc gcacgcgctc gccctcgtcg gctgcgcga cgtgcgcgac tacaaggctc 2340
 cgtcggctga cagcggcagg ctcgggacgt cgagccctt caacatcaag gtgcgatcgc 2400
 tcacgctcgc caacttcacc cgcgacgagg tcgcaacact ctacgctcag cacacggccc 2460
 agaccggta ggtcttcgg ccggacgcgg tggaccgcgc ctgcgagctc acccaggccc 2520
 agccgtggct cgccaaacgc ctcgcgcggc agctcgtcga ggtcctcgtc aaggaccgcg 2580
 cgcaacccat caegtctcgc aacgtcgatc gcgccaagga aatcctcgc gagcggcagg 2640
 acacacacct cgacagcctg gtggatcggc tgcgcgagcc gcgcacccgc gcggtgatcg 2700
 agccgatgct cgcgcgcacc gcgttgccga gcgtgcccc cgacgacctt cgttctcgca 2760
 tgcacctcgg cctcgtgcgc atgaccgcgg agggcgccct cgacgtcgcc aaccctatct 2820
 accgcgagat catcgtcgc gagctcgcgt tcccgatccg cgcctcactc cccagatca 2880
 aggccacgtg gctcacgcag gacggccgccc tcgacgcgga ccgctgctc gacgccttcc 2940
 tctcctctcg gcgcacgac ggcgagcgc tctcggcgcc cgcgcctac catgagatcg 3000
 cccgcacct cgtggatgat geettectc accgcgtggt gaacggcggt ggcaccgctc 3060
 agcgcgagta cgcacatcgg ccggggcagga tggatctctg cgttcgttac gggggcgaga 3120
 cgtcgcgat cgagctcaag gtctggcgag aoggcgcgcc cgatcccgtc gccgaggggc 3180
 tcgcccagct cgacgagtac ctggccggcc tgggctcga tcgaggatgg ctcatctct 3240
 tcgaccagcg ctcgggacag ccccccacg ccgagcgac gcgcgcgag cgcgcgctct 3300
 ccccgccgg ccgcgaggtc gccgtcattc gcgcctgagg gagctcgccg cgcggcgagc 3360
 gccctccaag agggccgggc caccctcgac agcgtctcta ctctccgag gccgcgcgc 3420
 ccccgccccc ggcgcgcgc gccgcgcgc gctccagctc gcagcgacc accaggacct 3480
 cgcacccgc gagctccggc cgtccacga gcgcgtgcgc gccgcgcgc accgcctga 3540
 gcaogtctcc cagcgcggc ttcagcgcgc ccagcgtcgc ggcgttcgccc ccgagcgca 3600
 ggtcggtcac gacgcgcgcc acgctcgcgc cgagctcgtt cttgcgcttg ttgaccgcg 3660
 ccatcgcgc cgccgcaga tccaggagcc ccggatccga aggcgcgcg accgcgcga 3720
 aatccgcgc tgaaggccac ttcgcccggg ggatcgaggt atcgcccgtc tctcccgct 3780
 acacccagcg ccagacctcg tcggtgatgt acggcaggac cggcgcgaa agccgcagca 3840
 gcacccagag cccgagcgc agcgcgcga ccgcgcgag gccgcgcgc tcccgcgcgc 3900
 cgcctcgcgc gcgcgcgcgc gccttcgga gctccaggta ggcgtcggtg aaccagcgcc 3960
 agaagaaatc ctcggctcgc tcgagcgcc cgcggaactc gtgctcgtcg aacgagcgcg 4020
 tcgctcgtc caccacggcc gacagcttgt gcaagagcgc ccggtcgagc tctcggaga 4080
 tcgggtggac ctcgcgcgc tggctgagca cgtacttgct cgcgttcag atcttcgtga 4140
 cgagccgctt gcgatcttc agcacttet cgtogaacgc cgtgtccgtg ccgagccgcg 4200
 cgtcgcga ccagtagcgg accgcgtcgc aagaatacgt gtgcgagcagg tgcacggcg 4260
 tgacgacgtt gcccttgctc ttcgacatct tcttgcatc cggatcgagg atccaccgcg 4320
 agatcgagac gtggtgccac gggaccgacg actcgtgcag catcgcttc gcgatcgtgt 4380
 agaacgccc cgtcctgatg atgctcgtgg cctgcggccc cagatcgccc gggaagagcc 4440

ggcgcgtggcg cgcgcgagtcg tccccccagt gagagctgat ctgcggcggtg agcgagctcg 4500
 tgaaccacgt gtcgaagacg tcggactcgg cgggtgaagcc gccgggctgg tcccgcgtcg 4560
 acgcctcgta cccgggcggc acgtcgaccg tcgggtcgac cgggagcacc tcgcgcgtcg 4620
 5 cgagcagcgg cccgctgtga tccgggttgc cctcggcgtc gaggcgatac cagaccggga 4680
 actgcacgcc gaaataccgc tggcggttga tgcaccagtc accctggagc cctcgggtcc 4740
 agttgcggta cgggaggcgc atgaaatccg ggtgccactt gatcttctcg ccgtattcga 4800
 10 ggagctcggc cttcttctcg gcgagccgga cgaaccactg ccgcgtgggc acgaactcga 4860
 gcggtcggtc gcccgcgtcg tagaacttca ccgcgcgtc gatcggcctc ggctcgcccc 4920
 gcagcgcccg ccccgggcgc gggcgcgcgc cgtgtcctc gcggcgagc agctcgacca 4980
 ccgcgcgcgc cgcctgctc accccctgc cctggagcgc cgcatacgc gcgttgccg 5040
 15 cggcggggtc ggggtctcc caccgcgcct cgcgaacgt caccggcagg acacggccgt 5100
 tcttgccgag catctgcgg agcgggagct tctgtcccg ccaccagatc acgtcggtcg 5160
 cgtcgccgaa ggtacagacc atcaggatgc ccgtgccctt ctcgcgatcc acgagcgggc 5220
 tcgggaagat cggcaccggc gcgcggaaga tcggggtgag ccgcgtcttg ccgaagaggt 5280
 20 gctgataccg cgggtcctcc ggggtgcgcgc tgacgcgcgc gcaggccgcgc agcagctccg 5340
 ggcgcgtcgt ggcgatgacg agctcctcgc ccgtccctc caccgcgaac gcgatgtcgt 5400
 ggaacgcgcc cgattgcggg cgatcctcga cctccgcctg ggagaccgcgc gtctggaaat 5460
 25 cgacgtccca catcgtcggc gcgaagaccg agtagaggtg gcccttctcg tggagatcca 5520
 ggaacgcacg ctgcgcgcgc ctgcggcagc gatcatcgat ggtggcgtag tcgttccgc 5580
 agtcgaccga gaggccacc cggcggaaga gcgccttgaa gacctgctc tcctcgcgc 5640
 tgacctgtg gcagagctcg atgaagtgg gccgcgacac gatgcgcgcgc ggctcctct 5700
 30 tgatcgtctc cggcgccggc tcgggcaagg tcaggccgcgc ctcgtaaggc gtgcgcacgt 5760
 cgggtcggaac gtggaagtag ttctgcacgc gccgtcgggt gggcaggccgc ttgtcgtccc 5820
 agcccatcgc gtgaagatg ttgaagccgc gcacccggcg ctggcggaag acgacgtccg 5880
 tgtgcgtgta gctgaagacg tggccgatgt gcagcgagcc cagggcggtc ggcggcgggg 5940
 35 tgtcgacgac gaaggtctcc tcgcgggggc gcgacgggtc gtatcggtac gtcgcgtcg 6000
 cctcccacag gtcggccagg cgcagctcgc cgacgggcga gtcgaagtgc tcgggagcg 6060
 tcgcgggatc gatggagcgc aacgtctct tgcgtcgcac gtcgtcacct gcagaacaga 6120
 40 ccccgagga accgcccgcg gggcgggcat cctacgtcgt ccccggggtg ccgtcgaag 6180
 cgcgcgcgc cgcgcggcg gcgacccgc atcgcaccc cgcacccgc agagcccggc 6240
 ggctccgcgc gcgcgcgcgc gccgtccgtg gagccgagag gagaggccgc cgcgcaggtc 6300
 gtggaggacg ccggcggcgc cgcgcgggag atcgcggaga ggcggggcgc tcgatcgcgc 6360
 45 cgaggccggg ggctcagtc tagcgtcga cgtggacgtg cttgcgggtg acccgagct 6420
 cgcgcgggc gagctcgcgc acggacgaga ccacccgac caggccgcag atgaagacgt 6480
 ggcgcgcgc atctcgcgc ttctcgcgc gctcccggtg gagctcgggc acgtcgcct 6540
 gcacgtagcc ggcgcggcgc gccacgacg ggccgcgcgc cgagagcgtg atctcgtagc 6600
 50 ggatccggtc ggatccgcgc gcgagcgcct cgagctcgc gcggtagatg acgtcctcct 6660
 cgaagcgcgc gccgaacagg atccacaggt ggggcgcgcgc cagccccgc gcgagggag 6720
 cgcgcagcat gctccgagc ggcgtgatgc cgggtgcggg cgcgacgaac aaggaggcg 6780
 cggaateccc gggatcgcgc gtgaagagcc cgtgcgggc gatggcgcg agcgtggcg 6840
 55 cgggtcgcgc ccggtgcagg tgcctcgcgc ccgcgcgcgc ctgcacgagc gtgaccgcga 6900
 gatcgaagc gggcgagccg tcggcgcggc atgcgatgga gtaggcgcgc ttcacctgc 6960
 cgcgcgggag cgggaggacg aggttgaccc actggccgc ctcgaacaga aacgacctc 7020
 60 cgtcggcggc ctcgaacgag agctcgcgc cgaagggtg gaggggcggc gcggcgacga 7080
 ggcgggcttc gaacggttc gctggatca tggtcgggc ccggcggggc tcggctgca 7140
 ggcgcgcgc gtggcgaggt cttaccgcgc cctgcgcgc ggcaccaatc cgatcgcgc 7200
 ggggaaggcg ccgcgcgag gcgcgcaat gcgggaatca cgggttcgc ccccgtcgc 7260
 65 cgcgcggagc cgcggcggc ccgcgcgcgc gcgctcgcgc ggggagcgc cgcgggctct 7320

accgcacgcc catgcgccgg cgctgcggga tgttcaccgc cggccgggag cgtacctggt 7380
 tggggagcgc gcgcgycggg cggggatccc ggtgcgcggg cttctgcgcg gggagctgcc 7440
 ctcgctgagc cgggcgctgg tcgggcgact tggccgagcc cagcgcgaga tcggagacgg 7500 5
 ggagatgcgc gcgtcgtgc atagaatcct ccatggaato ggtcatcaac acatcgggaa 7560
 gagcaccag gctgaaagaa accttcgaag aaccggctct catacaccct ccattcatcg 7620
 tgcgaccccg gattcaggac ggatcgaacc cgcgaggagc gctggtctct tgggcctctc 7680
 cctgctcgtc cgaccggcgc cctctcgacg caactccgcc gttcgtcggg acgggacggg 7740 10
 ccgcctcgcc gcacgctccc cgtcagagcg actcagcgtc tcgacgtcag gagagatgac 7800
 gactcggccc gtgcgcgcc gacccttcgc gctcgggtgt tcgagcgcgc ggccagcgag 7860
 cgagggggcga tcgccaggag atcacgaatc tcccggccat cggcctccag cgcctcgggc 7920
 tcgttcgctc gtgcccccgc tcgtccccgc cgcgcgcacg acgcgagctc gcgcggggaa 7980 15
 ccgcggggcg ctgtcgtggc tgtgatgag cgacgataca ggggggacgc cgtgcctacc 8040
 tgggcaacag gcgtcctatc tctaccacgg cgagcactac ggtgagtgtc gccatgagta 8100
 ggccccctgag ggtccgcgcg acggagcgtg gtgtcagcga gagatgcgca tgggtggacgc 8160 20
 gggtacgcg tcgagaggga cactagcact cgacctcgat cctgcaccag actttttgtc 8220
 ggggagggtt gccctccgcg tggccgctgg ccgttggccg ctcgcgcgtg gccgtcgcgc 8280
 gctggccgct ggccgctcgc cgttggccgc tggccatgtg cgacgtgagc tcgagcagcc 8340
 cgcggctgac ggacagaccc cggagtccat cgagccggtg atgccgaacc cgcgaagcga 8400 25
 aaaaacgtat ccgttcggca ggtcgtggcc taccatgcaa gctgctcgat gcgctgacag 8460
 gcttcttcga gatccctcgc ggtctttgag aagcaaaacc gcataagcg actccctgc 8520
 gtccttcaa agaaggcgtc gcctggcagc cccgccacc cggctctcgt gagcaagtag 8580
 atggctcgtc ctcgacctgt cctcccggtt aggcgagaca catccgccag cacgtagtag 8640 30
 gtcctcgtgc gcacgcaggg tggcaagccc gctttctcca gcgccgaca gaaccggtct 8700
 cgttcctgtt catatocctg ggcaagcccc gtgtaaaacg agcgagggaag gccgcggatc 8760
 cgggcagcga ctcctatgct cagcggcgtc ggccgcgaga cctacagcag gtcgtcctatg 8820 35
 gctccaatgg ccttcgcccc cctggcctcg gccacgtgt agccgatccg ccctcctgtg 8880
 atgtgaagg tcttcgagta gccgcctatc gtgatcgtac gctcggacat gccgggaagg 8940
 gaggcgacgc tgacgtgctc acggccgtcg aagataaagt actcgtaaat ttcgtccgtg 9000
 atcccatga ggtcatggtg gcaggcgaga tcggcgatct gttccagctc cattcgcccg 9060 40
 aacaccttc cggaaggatt tcagggagag ttcaccacga tcgccttggg cttcgggggtg 9120
 atcgcgcgtc ccagctcgtc gccgtcgaca ttccagctca gggatcgcgc cgtcacatac 9180
 cgcggaacag cctcgacggc gaggatagcc tgggcgtgat aggcataaaa cggctcgaag 9240
 agcagcactt cgtccccagg attgagcaag gccatgcaag tggcctgaaa ggccctgtc 9300 45
 gctccggcgc tcaccgtgat gtcagctctc ggatccgcgc cgtatgccatt atggcgagcc 9360
 agcttcgcgc cgtatcgcat gcgcagctcc acgatgccgt cgaagcgca atagtattg 9420
 ccccccgat ccatcgctc cttaccgct tgaaggatca ccgaaggaaac tggggatca 9480 50
 cagacgcctt gggacatatt gatcccatgg accttggcgc acgccagggt catggtacgg 9540
 atatcggaat gggcgaggcg agccgcacga tccctcggtg gactcttcct cagcgtgctc 9600
 ctgcttctgt tctgcggctc tgcattggtt cttcgggttg gcttgtcagc tcgacgcgcc 9660
 catgcagcgg cgcagcccta gcggccgcag gtctgtccac acttctttga tgaaagcgag 9720 55
 acattcggct ttcgtgcctt gtttgccgc agccctccag cccccaggta cgggcttgtc 9780
 ggccggccag atcgagtaet gctcttcgcc gttaccacg acctggcaac gcgtcttgtc 9840
 ttcgtcgtcc cgattcatga tttctctcgc ccttcgtcag cgtgcgcga gcatgaaacg 9900
 aatcgtcat cggcgacag gcgcgcgcgc gctgcgcgga ggactccca cgcctccctc 9960 60
 acggcaacct catcgtccg gatgttccc atggcgactc ggatcgtgta cctgcctgtg 10020
 agacgggtat gggacaaaaa taccctgcc gacttgttga cctcgtccag cagcgcctc 10080
 ttgagcgat cgagctcgcg ttcgatcgac tctctctctg cctcgtccgc cgaccgcatg 10140
 atgcaagcga gcgcggaggg cctcatcgca aagcagaccg tactgaacgg cgtcggcgcg 10200 65

aggcgetccc aatcgggata gggtccacc cactgggoga gctgctgccc caateggagg 10260
 tgctcccgga tccggggcgc cagcccttca tgcgccgaagt agcgcacgat catccagagc 10320
 ttcagcgctc ggaagcgccg accgagctgg ataccccagt ccatgtaatt cgtgacgtcg 10380
 5 cccctcgggtgc ggaggtatcc gggeaccaga ctgaacgcgc gcttcagtcg gtcggcgtea 10440
 cgcacgtaga gcacgctgca atccatgggg gtgaacagcc acttgtaggg gttcactacc 10500
 agcgagtccg cccctcgcga gcccgcgagc acgtccctgt gctcggggac gatcggggcc 10560
 10 atcccccggt aggcgcgctc cactggaagc catagcccgt gctcccgga aacgctgacg 10620
 atggcgggga tggggctgac gctcgtcgtg gacgtcgtgc ccaccgtcgc cgcgacgcag 10680
 aagggtcgga ggccggcccc gaggtccctc acgacggcgg cgcgcagcgc ctccggggacc 10740
 atcggaagg ccggtaccgt ggggatcttc cgcacccct cctgcccgat gccgagggtg 10800
 15 atggctgcct tctcgatgga tgagtgcgcc tgetccgacg cgtagagtcg catgcgccgc 10860
 tgctcccgca tgcgccggag ccggtgggtc ggtcgggccc agtcgcgcgc gcccgcgatc 10920
 gcgaccatgc tggcggtcga cgcggtgtcc atgatcgcgc cgtgcaagcc gccgtcgaga 10980
 tccagcatct gacgcagcca ggagaggacg agctccctga gctcgggtggc cgcggcgac 11040
 20 gtgcgccata gcacacgtt gacgttgagg cagcgcgcga gcagctcgcc gaggatccca 11100
 ggaacagacg ccgtgttcgc gaaatacgcg aagaatcgcg gatgattcca gtgcgtgac 11160
 cccggcagaa tgatctgctc gaaatcgggt agcacggcgt ccatcggtc cggctcgacg 11220
 25 ggcggggctg gggccagcct gcccttcacg tcgcgggggc ggatcgcggg aaagacgggg 11280
 tatcgatccg ggtggccgag gtaatcgcc gcccaatcga tgattctcat accgatccgg 11340
 cggaaactct ccagatccat gtcccgagc cgttcttcc gcgggtcgct cagtcaccc 11400
 tctcgcctc gccaggacag gatcctcgag gtccctggc tccggcggtg gaaagcgctc 11460
 30 cttgaacgtg aaggcccacg gggctcggtc gtacgcgcgc aggtgctcga gccgatccctg 11520
 cccctcgcgg accgacggga tgtgcccgcg cgggacccac cacagcacga ggtaatgcgg 11580
 ctcgagatgc tcgaaccacc gagcgcgctg tcgcaggaaac gcggcatgat ccgcggtgta 11640
 ggtgaaggcg aacaggtgct cgatggaggt ccataccgac agggtcacga ggagccgctg 11700
 35 gtccgggtac ggacggatgg acacagagtt cccctcggcc gtctgcaggc gccacacgaa 11760
 cccctcgcctc cgatcgccca gatggttgat atggctcagc ccctggacga agccctccat 11820
 gatcggtacc tccagcggag cgcgaataca tgcgaagtgt tattgcgcga tgtggtgccc 11880
 40 atgtcccgac atgtcgttt ccatctccag ctcccgctca ccaatcccag cgtgctccg 11940
 gggagctcat cagggcagac gcgacatcga tcccgaaagt ccgcgcgcac cctcgcagca 12000
 aggggcctg gaccgcttcg gcgacggatc ggctgcctc cggcaagacc tcggagacaa 12060
 agaagaaccg cctcgtggaa gggacaatct tgcccgcctc cgcctggcgc catacgaagt 12120
 45 gcctcgtcac cagtccctcc gcgtcgccat acccgacctc gccggcgccg accgccacgc 12180
 tgccgcctga gccgagctcc acgaacgcct caccgcctcg cgagatctcg aggcgaacgt 12240
 ccggggccagc cagatcgccg aggtcccaag cgcgcagcgg gacggcgaaac cgcagcgaca 12300
 ggaggttgta aaaatcgacg aatgcgttga tgtgcggcag ctctccacca ccgaggacce 12360
 50 gcttcgcccag cgcctcgatc gagctcgga atttcttgcc agagaccccc actcgttcca 12420
 tcgcctcgcg ccaggcagcc acgtcggtg gcgactgggc gtttctggtg ccccgctcc 12480
 gtgcagctc ctccctgacc ttccggagct cctccagcac ggccggcgcc tctcgggcgt 12540
 55 tgtccaggcc ttcccgctac ccggtgacca agatcatccc aggaaacgac tcccgatcc 12600
 gcggtcgac gatgaatgcc atgtgcctcc tgcccctcga gagcgatcgc ctcgatcgac 12660
 accaggctgt ggatgcatga gccgggcctg gcggacgcag gaccccgcta ctcatggctc 12720
 ttcgtggccg atgaacaggt cctccacccg tcgatcgtgc tcggtgcccc gatccgtcca 12780
 60 gtccaccccg ccggcgaccg cgatgtttgc acccgagacg tacgaggcgc ggtcggaggc 12840
 gaggaaacgc acagcatctg cgacctcgtt ggcgcgcgcc aggcggccca tggggacgcg 12900
 ccgctccatc cactcttct gcgcgggcgg aaggtatccg ttgtcgatga gccctggaga 12960
 cacacagttg accaggatcc catgaggcgc ctccctcgtg gccaggctgc gcgtgaggat 13020
 65 gagcacgcgc gtcttcgcga tcgagtaacg cgcacgctc ggccgcgcgc ggatcgcgta 13080

cgtggggctc aacccgatat tgatgatccg gccgctcttt cgtggcgca tgcgcgccac 13140
 ggccgcgcga cagaggtaat gaacgctgct cagggttctg tccatgacgt tgcgccattc 13200
 gtgcgtcggtc atcgccgcaa ggggcttgaa gaagaagtcg cccacgttat tgacgaggat 13260
 gtgatgggg cccagctgcg cctcgacgct ggagaagagc tccgcggccg cgttggggcg 13320
 ggtgacgtcg gcctgcacca ccatggttcg tcgcccagc gcgcggatct cggccgcgct 13380
 ctgctcgccc gcctccttgt tcgaatggta attgacggcg acgtccgcgc cttgtctcgc 13440
 gaggcgcagc gcgatcgct tgccaattcc gcgcgagcta cgggtgacca gggcgacgcg 13500
 cccggcgagc tccagcgatc gcgcctgtgg cagggcggga gcagcctcct ggtggagctc 13560
 gacgtcgacg gggagctcca cgtggtagct cgtctctcgc ggagccgcgc agtacctctc 13620
 gtagaacgcc tcgaggacgg gctcttcgcg tcgcatgatg tccgcgtggg attcggcgct 13680
 gcgccacgga tagagcacca ggatctcgtc ggggcgcacg gtgctctgaa agaatcggc 13740
 atgaccgcac caccgccggg gttcgtgggc tcctggggcg agcagatcat ccattttttg 13800
 catgatctgc gtggcctcgc cctccatgcc ggggtttgat gcgcatcgct ccataacgag 13860
 gatcatgtct tggtcctgt tcgtcatcgc cgttctcgc tgggggggct gccgcgcctc 13920
 tcgagggcgc gcccttgta ttggccgcgg atggtctggg tagcgtcgc gagctttcgc 13980
 ttgtggggcg cgttcaggct tgcgcctga ttgacgaacc gctcgagac gaatgcgtgc 14040
 gattcatatg cgttcgcgag cgcgccacgg tagagacggg ggtacgagc gagcttgggc 14100
 ggaagcgcgc tcaaggcggg tcggatggat ttgaggggtc gaatcaagcg ggcgtgctct 14160
 tcgaagaaga gaccgctgaa tccatcctgg acgaacggcg gggccatggg aggtcgcacg 14220
 atgttgttgt aatcatcatt cgagaaatcg ccggtaaacc ggaagatgc cgcggttgc 14280
 caccagagct cgttgagat cgcgagcgc tctcgccgc tcggaacgct acggttggcg 14340
 agcgcctcgc acagccacct gtgcgccatg agcagcgct ggaccagcac gaagaacgcg 14400
 tgatgcccc aatccagcg gccagcgcg ccgtccggag cgcgcgcctc cgcggagcgc 14460
 gcagccagct cggggatgcc attocatgat tttggccttc gcttgcggga gaactggtgg 14520
 aggatgtcct cgtggagc acagagatgc cccatttcca tgggttgag ggaagtacct 14580
 ttcaggcttt cgcggatcat tcggtaatat gcgacgatca ccgcttcgca gtacgtctcg 14640
 agcgggtctg gggcggagac ccggtgcaca tggaaatagg cgtcgtactc cgcctcccg 14700
 tccgacagc tccccggat gtcggatct gccgcttgc cctgccatcg ctcgaggatg 14760
 gggacggcga gggggcgag atgctcggcg agcaccgcca ggtcgccctc cggcgcatgc 14820
 gccaacaaga cctcgaaggc ctcggctcgc accttgagg tctcgcccat tccgacggcc 14880
 tcgatggctt gggcgagcg tagatggatg gtatatttag ccatgatttg cccgaagatt 14940
 gcgctcgt cgacagatct ttcgcgagc ggaacgccat ttcactgt ctggtctca 15000
 atattgaatt gagccctggc gactgccata ggcccagtc ctcgacacag tgtaaggagc 15060
 ggcccgatgc tttctcctt tttagtctc caccgaatac ttctgttggg cgccaaagat 15120
 cccttgccga gactgtccg cgagatgtcg tgtgcgaagc gtccgcacgt ccagcgggccc 15180
 catgcgttgc tagagcataa aacggttcga tgccgtgctg agagggagac gcgaggagcc 15240
 tccctttggg acggatgagg aatttcgtga ccgaaatgc ggcaggaca ggcgcgcaga 15300
 agcggcgcct cgtggggaa ccatgggtta cgaagacatt gatgataatg tcgacgcaat 15360
 cgcaatcgtc gcgatgagcg gccgttccc cggcgcgaga aacgtcgagg agctgtggca 15420
 gaagctcgc gctggcgtgg aatgcgtcgt caccctcaca gaggcggagg cgtcgcgcgc 15480
 gggggtgagc cgcgagatgc tcgcgaatcc cagctacgtg cgcagaggcg cgcgctcga 15540
 cggcgtggag ctcttcgacg cctcgttctt cgggttcagc ccgaggagg cagagagcat 15600
 ggatccgcag cagcgcctct tcctggaggt cgcctgggag gccctcgagc gcgcgggta 15660
 cgaccccgat gccattccg ggcctatcgg cgtcttcgcg ggcagcgccc cgagcggcta 15720
 ccaactcctg gcgcagtcgc acccgagat cctaggcgcc ctcggccact accaactgac 15780
 gctgaacaac gacaaggatt atctcaccac acacgcctcg tacaagctca atctgcgggg 15840
 cccgagcgtg tgcgtgcaga cgtctcgtc gacctcgtc gtggcgtgg tcatggcctg 15900
 ccagagcctg ctcaaccacg agtgcgacat ggcgtcgcg ggtggcgtgg ggatccatgc 15960

5

10

15

20

25

30

35

40

45

50

55

60

65

gcatcagcgg aggggctatc tgtatcagga gaacggcacc tcttcgcccg atgggcattg 16020
 ccgcgccttc gatgtggccg ccaagggcac cgtgggcggc agtggcatag gcacgtcgt 16080
 cctgaagcgg ctgcgcgacg cgtcgcgcga cggcgaccac gtgcacgcgg tgattcgagg 16140
 5 agcggcgatc aacaacgacg gctcgcgcaa gatcgggttac accgcgcgca gcgtgcaggg 16200
 gcaggccgag gtgatcgga tggcccaggc gctcgcggc gtggagcgg atgacatcag 16260
 ctacatcgag gcgcacggca cggggacgcc gctcggcgat cccatcgaga tcgcagccct 16320
 10 caccgcgctg ttcggggcga agaccgcagc aaggcagttc tgcgccatcg gctcgcctca 16380
 gaccaacctc ggccacctcg atgcgcgcgc gggcgctgcc tcgctgatca aaacggtcat 16440
 ggccctcgag caccgcgagc tgcgcccgag cctgcacttc gagcgtccga atccgaagct 16500
 cgagctggag agcagccctt tctacgtcaa caccgcctc actccgtggc acgcggcacg 16560
 15 aggtccgcgc cgcgctggcg tcagctcgtt cggcaccacg cgcacgtgg 16620
 cctcgaagaa gctcgggcc cgcctccgag cggccctcg cggcgctggc agctcctcac 16680
 cctcgcggct cgtcgcgag cgggctcgc gcgggccacg gccgacatga tcgagcact 16740
 cgatcgccac tccggcacat cgatcgccga tgtaacgtac acgagccacg tggggcgccg 16800
 20 ggcctggccc ttcggcgag cggctcgcgc cgagagcgcc gcggatctcc gcgcgcgct 16860
 cgcgagcgag ggtcgcgcgc gctcgcctc gtcacgcag gcggcgagg agaggcccg 16920
 cgtctctctg ttcgccggtc agggagcgca gcacctctc atggcgcggg agctgtacga 16980
 25 ggtcgagcgc atcttcggc agtcctcga ccgctgcgc gagctcctgc gcggcccgct 17040
 cggcctcgat ctgcggcagg tctctaccc cgcgagggg cagcgcgacg acgcgagca 17100
 ggagctcgg aggacgcga tcgccagcc cgcgctgttc gccatcgagc tctcgcctgc 17160
 caagctgtg atggcctgg ggatcgtccc ccaggcgatg atcgccaca gcgtcggcg 17220
 30 gttcgcgcgc gctgtgtctg cgggcatctt ccgcgaagag gacgcctcc gcctcgtcgc 17280
 cgagcggggc cgcctgatgc aacagatgcc gccggcgcg atgctggcg tgccctcgc 17340
 ggagcccgag ctgcgccct acctcagca cgacatctcg ctgcggcgga tcaacggtcc 17400
 ggctctctcg gtggctcgtg ggcgagtcga ggccatcgac gcgctcgcgg ccgagctctt 17460
 35 ggaccacggg ctctcgtgcc ggogactcca caccggcac gccttcact cgaagatgat 17520
 ggcccccgtc gttgacgct ttaccgatg cgtgtcgcgc gtcgagcgcc gccgcgcgc 17580
 aggcacttc ctctcgacc tgacggcgcg ctggatctcc ccgaagcag cgaccatccc 17640
 40 cgcatactgg gcccgcgagc tcgtggagcc ggtgcgcttc gccaggcg tgaggcagct 17700
 gctgtccgag tcgacgtggc tctggctcga gctgggtccg gccagaccc tgagcccgct 17760
 cgtacggcag caggcccgcg cggatggcg ccagggtggtc gtcgcctcgc tgcgcgcgc 17820
 gaaggacgc ggccgcgacc acctcgcggt catcgaggcg ctgcggcggt tctggagcgc 17880
 45 tgggtgggac gtcgactgga agcgtttca cgaggcgag gcgcggcggc ggggtgctgt 17940
 accgacctac cccttcgagc ggcaacgata ctgggcctct ccgcgccaca cgagcgctcc 18000
 gccggaagcg ataataagc cgtcctcgc gaagaacca aacgtcgcg attggttctt 18060
 cctccctgcc tggcgcgct cggatcctcc ggtctcgttc gacgcgcagg cggtgaccac 18120
 50 gggcgctct acgtggctcg tctcatcg ggacgagggc ctgcggcgcg cgctgggtgga 18180
 gggcctcgc cgccgggggc acgaggtcgt cgcggtggtc acgggtgaga ggttcgagca 18240
 gacgggcacg cagcgtaca cgatcgatcc cgcgcgaat ggcgatgtt cgtccctctt 18300
 55 cgcgcggctc gaaatcgaag ggcgcatgcc ggaccggatc gtccatgct tctgcagtc 18360
 gcctcgggac ggccgcgcga tcgagcgcg agcgcgctg gagatcgagc gcaggctggg 18420
 ctctgatagc ctctctctc tcgccaggt gatcgccgca caaaggcacc cgaagcgcct 18480
 gatgctcggc gtgatcacga cccggcgca ctcgctcacc ggaaccgaga tcacgagcc 18540
 60 cctgcgcgct ctgggtgctg gccctgcgc cgtcatcccg caagaatac cccatgtctc 18600
 gtgcgggaac atcgatatcg atctccggg cgaaggcggg cgcgcggaga tcgcggcgcg 18660
 cctgatcgcc gatctggagc gagagtcgc cgaactcggtg gtggcctacc gcggcgccg 18720
 gcgctgggtc gagagcatag agctaccga tgcggcgcg cggtcagctg gcgcgcgcc 18780
 65 gcgcctcgc cagcgcgggg cgtacctcat taccggcggc ctggggggca tcggcctcgt 18840

DE 198 46 493 A 1

ggctgcagag ctcttgcccc gagaggcgca cgcacggctg atcctgggtg ggcggacagg	18900	
cctgccagcg cggcaggggt gggacgactg gctcgcggcg caccgcgagg ggcacgcgac	18960	
gagccgaaag atcctccgga tccgcgcgct cgaggaggcc ggcgcgagg tgaagatcgc	19020	5
cgcgcccgac gtctccgatt tcaatcgat ggcgagcgtc atcgaggagg cccggacgcg	19080	
cttcggccgc atcgacggcg tcattcactc cgcgggcacg gcgagtggag gcatgatcca	19140	
gctcaggacg ccgatggcgg cttggcgcgat gatggcgccg aaggctggcg gcacgctcgt	19200	
gctcgatgcg ctccctccggg acgagcgtcc cgacttctc ctgatctgct cgtcgttgge	19260	10
ctcgetggtc ggcggcgcca cccagatcga ttaactgcgc gccaacgcct tctcgaagc	19320	
ctacgcgcag agccgcgagg ggcaggagg atgcgcgtc atctcggtgc aatgggacac	19380	
gtggagtgc gtcgggatgg cgggtggactt caagctccc gccgatctcc aagaggggag	19440	15
ccgcgagagc ctgaagcggg gcacagctc gagcgagggc gccgaggtgc tcggccgcat	19500	
cttgagcgca ggcagtgagc gcccgctggc gatttgacg tcggatctac cagcgtacaa	19560	
gcagtctgtc acgacacgcc gatcgacga tgagcaaac cccgcgcgcc ggcgatgca	19620	
ctcgcgcaca acgacacagg gacccatgt cgtcccgag accgagaccg aacggcgcat	19680	20
cgccgcgacg tggcaggatc tcttcggcct cgagcaggta ggcgcaaac acgatttct	19740	
ccagctgggc ggccattcgc tgttgccac gcaggctctg tctcgcgtcc tgcagacct	19800	
caagggtggg atctcgttgc cgcagttctt cgatgcgcg acggtcgcag ggcttctcg	19860	
cctggctgac gcagcacggg ccgaaggcgc cggaccgtc gcgcggcaa tcggcgtgt	19920	25
cgagcgagac gcctaccgaa tcaagcgcc cgcggccgaa caggccgcc gcaccaagc	19980	
gtaacaagaa ggggatcgag tcattggaacc cgtcggcggc gtggacatga atcagccgc	20040	
aaagcagcag gagacctgc tctcccgac ctcttcgcg cagcgcggc tctggttct	20100	30
cgaccagctc gagccgggga gcgcgtctc caacatgcc gcctccttc ggacgcgcg	20160	
gccgtacgac gtcgactcgc tcgtgcgcag cgtgaacgag atcgtgcggc gccacgagtc	20220	
gctgcgcagc accgtcgatg tcacgatgg cgaaccgtg caggtgatcg cccctcgt	20280	
gcgcatcgag gtgcccgtcg tggacctgag cgagatcgac gacccggagc gagaggcgga	20340	35
ggcccgcgcg ctcatggcgg aggagagcgc ccgcccttc gatctcagc gagggccgct	20400	
gctccgagcc aagctgctcc ggctcggcga ggcgatcac gtgctgatct tgacgatgca	20460	
tcatactgc tccgacggct ggtcgatgga cgtgctgttc aaggagcttt ccacgtcta	20520	
cgcgccttc cagcagggcc gcccgctgc gctcccgag ctgcccgttc aatacgcca	20580	40
cttcgggtg tggcagcggg agctgctcca gggcgaagt ctggaatgc acctcgggta	20640	
ctggagagag cacctccgcg ggcgcgcac cgtgctggag cttccgatgg accggcccc	20700	
gccgcggcg cagacgttc ggggtccca gcgcgcgttc cgactccac tctccctgca	20760	45
acaggcggtg caggcgtca gccgcagga aggcgcgacc ccttcctga cgtgctgac	20820	
ggcgcttcagc gtgctgctct cgcgttatgc gcggcagagc gatctgggtg ttggcagcc	20880	
catcgcaat cgcacccgag cagagctgga ggggctgac ggcttcttc tcaacatgt	20940	
ggcgtgctgc atcgacctc ggggcgaccc gagcttccgc gagctgctcg ggcgggtgc	21000	50
ggaggtgacg ttgggcgcct acgcgcacca ggacctgcc ttcgaacggc tgggtggagga	21060	
gctgtcacca gggcgagcc ccagccacag cccctgttc caggtgtcct tcacgttgca	21120	
gaacaccccg atggatgca cgaacagagc agacattgca tcgggtggcg cgcgcgtgt	21180	
ggaaatgaag gcggcgaaat tcgatctgat cctggagctc tcggaatgc cgcaagggt	21240	55
gctcggcacg ttcaggtaca acaccgacct gttcgacgc ggcaccatcg agcggtg	21300	
cggccacctg gaggtgctgc tctccagcgc cgtcgcggcg ccggatcgac ccattcgga	21360	
gctgccgtc atgggggccc aggcgcgag tcgggtattg gtggagtgga actccactgc	21420	60
cgcgtgtat cccgaggacc attgcatgca cgagctgttc gagcagcaag tggagcggtc	21480	
gcccaggcg acccggtgc tctccagca gcagacgttg acgtatcgag agctgaacat	21540	
gcgcgccaat cagctcgcgc atccctgcg gagcctgggc gtgggcccag aggtgcgcgt	21600	
cgggtgtat ctgaacggc caatcgagac ggtcgtggcg atcctcggcg tgcacagga	21660	65
tggcggggccc tacgtgcgc tcgatccgac gtacccagc gagcgctcg ggctcatgat	21720	

ggaggacgca ggcgcctcgg tgctgtcac gcaggcgctg ctcctctcga agctgccgcc 21780
 ccacggggat gcaacgctgg tacagctcga cgcgtgcac gaagcgctct ccaggctgcc 21840
 acaccatacc ccgcgagcgg gcgtcacgc ccagaacctc gcatacgtca tgtacacttc 21900
 5 cggctcgacc gggcgcccca agggcggtgt cgtcgagcac cggggcctct gcaacctgcc 21960
 caccgtgcag gccaaactct atggaatcgc gccggggcgac aggtcctccc agttcgcgcc 22020
 gctctgtctt gacacatcgt tctgcgagat cgcgtcgcg ttgtctctcg gagcgagcgt 22080
 10 ggtcatgggc acggcgagc agcttctccc gggacctccg ctggctcgagc tgctgaagaa 22140
 gcacgcgggc acggcgatgc tctggcccc tacgtgtct gcagcgctgc cagaacaaca 22200
 gagcgcgggc ttgcgcgtgc gcgtgtcac gatggcggtt gaggcggtgc cggcgagcgt 22260
 cgtcaagcgc tggaaggcac ccggacggcg cctgttcaac tctatggcc cgaccgagac 22320
 15 gaccatttgg gcaagctccg cagcggaact gtccgacgaa cggatcccg ccacggcgcc 22380
 tccgattgcc aatacgcaaa tctacgtgt cgacgaagcg ctcgagccgg tgcccatcgg 22440
 cgtgccgggc gagatcttca tcggcgcggt gggcgctgcc cggggatata acggcgctcc 22500
 ggacctgacg gccgagcgat tcgtaccgca ccccttcggg caaaccaaag gggcgcgcc 22560
 20 gtatcgagcc ggcgatcggg cgcgtgtggt gccggacgga aacctcgagt ttctcggtcg 22620
 aaacgacgag caggtgaagg tcccggtgt ccgcacgag ctggaggaga tccgcgggc 22680
 gttgtcaag caccggcggt tcgtcaagc cgtggcggtg gtgcgcgagg acacggcgcc 22740
 25 ggacaagcgg ctcgtcgcgt atgtcgtcgg acgcggagga gcgcgcgtga ccgcgcggga 22800
 gctgcgcag tccgtgagcg agcgattgcc tgcgacctg gtgcatcgt ccttcgtggc 22860
 gctcgacgcc ttgcccctga cgcgaatgg caaggtggac cgcgcgcgc tgccggagcc 22920
 cgagcagagc gccggcgggc aggaccacgt cgcgcgcgc aacgcgcgc agggaggagc 22980
 30 cgccaggatc tggcgagcgg tctccgggt cgaaagggtc ggcgtccacg acaactctt 23040
 cgagatcgcc ggcgactcga tctgagcat ccagatcgt gtgcgcgcgc agcaggcagg 23100
 gctgcgcctc accccgcgtc agatgttcca gcaccagacc atcgccgagc tttcgaccgt 23160
 ggctagagcc gtcgagcggt tccacgtcga gcaggaaeee gtgaccggtc ccgcgcgcgt 23220
 35 caacgcgggt cagcgtggt ggtcggagca ggaggcgcc gagccgacc acttcaacca 23280
 gtcgatcttc ctgaggtac gcgagcggt cgacgagagc gcgtcggagc aggcacgc 23340
 gcactgtatc gaccaccacg acgcgtccg gttgcgcctc gcgcgcgacg aacgcggcgc 23400
 40 ccaccaggtc ttgcgcgcgc cgggaggtc gacccattt cagcgcgctc acctcggggc 23460
 gctgcccagc gccgagcaga tctccgcat ggagaaggcc gcgagcgagg cgcaggcgag 23520
 cctcgatctg gccgcgggcc cggctcgtcc gcgcgtgctc ttgacctcg gcgaggtcgc 23580
 ccgcgaacgg ctgctcgtca tcgcccacca tattcggtc gacagcgtct cctggcggt 23640
 45 cctgctcgac gatctctttg gggcctatga gcaggcggc cgcggcgagg ccgtacgcct 23700
 gccgcccagg accacgtcgg tcaagcgtg gcccgagctg ctcaccgagc acgcgggctc 23760
 cgaggcgctc aaggcgagc tcggtactg gctcgactca tcgcgacgaa cggtagctcc 23820
 gctgcccgtg gatcgacggg ccggcgagga cgtgtggggc tcggcgcgcc acatcgtcgt 23880
 50 ctcgtcagc ccggagcaga cggagcagct cctgcgcgag gtgcccagc cgtaccgcac 23940
 acggatcgac gacgcgtcc tcaactcgtt cgcgcaggcc atcgtcgggt ggacgggctc 24000
 gccggcggtg ctcctcgacc tcgaggtcga cgggcgcgag gayctgcgc gcgtagacct 24060
 cagcgcgacg gtcggctggt ttacggccat gtaccgac ctaactcgcg tcgacgggc 24120
 55 ggatccgggt gaggcgtcga aatcgatcaa ggagcagct cgcgcgtgc caggcccgcg 24180
 gctcggtac ggcttgttgc gttacctcg gtcgatacc atcgccgagg tccgcgcgtt 24240
 gccgcaggcc gagctctgt tcaactacct cggccagctc gatcaggcga tccccaggc 24300
 60 tgacacgttc cggccgggc gcgagatca aggtcggag cgcagcccg gcgcccacg 24360
 cgccacctc atcgaggtga acgcgagcat cgccaatggg cgctgtacg ccacgtggac 24420
 gtacagcgag cgcgcgcac agcccgaaac catcgagcgc gtcggcgga gcttcgtcac 24480
 ggcgtccgc gcgtcatcg cgcactgcac cttgcccag gtcggcgga acacgccttc 24540
 65 cgacttcgac aaggtgcgc tcgcccagga gaccatcgat gctctcgac caatcgacgc 24600

gggccccggg cegtctgcga gggggagccg aatcgaagac gtctaccgcg tctcgccgct	24660	
ccaggagggc atcctgttcc acacgcteta cgccaccgat tacacggcgt atgtcgagca	24720	
gttccactgg acgctggagg gcgatttcga cgccaggcgg ttcaccgcg ccctccagga	24780	5
cgtegtcgct cggcatgccc coctgcgcac gtcgttcgccc tgggagcgcc tcgatgctcc	24840	
acttcagatc gtccgcacgg gcgcggtcct ccccgtcgag caccaggacc tacgcggcct	24900	
cgcccgaggag gaggagaccg cgcacatctc cegttagctc gaggcagagc gccagcgccg	24960	
gttcgatctg cgaaaggcgc ccctcatgcg cgccgggctg ctccggctcc gcaaggacgc	25020	10
ctggtgcctc gtcgagacca tccaccacct gatcctggac ggctggctga cacaatctt	25080	
gtctaaagaa gtgttcacgc tctacgagge gcaccgcgga caccgtgggc atctcgcgct	25140	
ggagctcgag cagcccgccc cctacggcga ttacatcggc tggctcgca agcaggacca	25200	15
ggtgcgcacc gcggccttct ggccggcgca gctcgagggc ttctccgcgc cgacgcgct	25260	
cgccgctgac cgcgctgtgc cgcacgacga cgccggcccc cggtttggtt ggccgcgcat	25320	
cgccctctcg ggcgacgacg cgcccggtct cgccgcttc gcgcgtcagc atcagctcac	25380	
gatgagcagc ctggtgcaag gcgcgtgggc gctgctcttg tcaagctaca gcggcgatcc	25440	20
cgacgtgctc ttccgtatga ccgtctcgcc cgctcgccgc ccgattcccc gtatcgagcg	25500	
catgaccggc ctcttcatca acaccattcc ggtgcgcgtg cgcgagcctg ccgacgcgtc	25560	
ggtgctcgcg tggctcaagg cgctccagga gcacgaggca gagctgctcg agcagcagca	25620	
cagcccgctg gtcgaggtcc agggccatag cgacgtgccc cgccgggaccc cgtctctcga	25680	25
gagcctcgtc gtgttcgaga actaccgggt gcaggtcctc ttccaggccc ctccggctga	25740	
ggggccgacg cgccgggagg agggcctccg catgatcgat gcgcagtata tcagtgatcc	25800	
accgtatccg ctgacggtcg tcgcccgtct ccatgggacg ctttatctca atattggcta	25860	30
cgagcgccgc cggttcgacg accaggccgt cgaacggatg atccggcagc tcacgacgct	25920	
gctccggggc ttcggtgcga ggcccagac gtcgggtccc gatctgcggt tgetgacggc	25980	
cgaggaggag cgcacccagc tccacgcgtg gaatgccag gcgcgcgct atccccaggg	26040	
ccattgcatg caegagctgt tcgagcagca agtggagcgg tcgcccaggg cgaccgcggt	26100	35
gctcctccag cagcagacgt tgacgtatcg agagctgaac atacgcgcca atcagctcgc	26160	
gcacacactg cggagcctcg gcgtgggccc agaagtgcgc gtgggcttgt gtctcgaacg	26220	
gtcgatcgag acggctcggt cgatcctcgg cgtgctcaag gcaggcgggg tctacgtgcc	26280	
gctcgacccc acgtacccca gcgagcgccct cgggctcatg atggaggacg cggcgccctc	26340	40
ggtgctgctc acgcagacgt cgtcctctc gaagctgccc cccacggggg atgcaacgct	26400	
ggtacagctc gacgcgctgc acgaagcgt ctccaggctg ccacaccata ccccgcgag	26460	
cggcgtcagc gcccagaacc tcgcatacgt catgtacact tccggctcga cggggcgccc	26520	45
caaggcgctg ctcgctcgagc acccgggcct gtgcaatctg cccaccgtgc aggccaaagt	26580	
ctatgcaatc gcgcgcgagc accggctcct ccagttcgcg ccgctctgct tcgacacatc	26640	
gttctcgag atcgcgctcg cgttgctctc gggagcgacg ctggtgatgg gcacggcgga	26700	
cgagctcctc cggggacctc cgtgggtcga gctgctgaaa aagcacgcgg tcacggcgat	26760	50
gctcctggcc ccttcggtgc tcgcagcgt gccagaacaa cagagcgcgg cgttgccgct	26820	
gcgcgtgctc gcgatggccg gcgaggcgtg cccggcgag ctcgtaagc gctggaaggc	26880	
accgggacgg cgctgttca actcctatgg ccgacccgag accaccattt gggcaagctc	26940	
cgcagcgagc ctgtccgacg aacggatccc gccatcggc cgtccgattg ccaatacga	27000	55
aatctacgtg ctcgacgaag cgctcgagcc ggtgcccatc ggctgcccgg gcgagatctt	27060	
catcgccggc gtggcgctcg cccggggata tcacggggcg ccggacctga cggccgagcg	27120	
attcgtaacc gaccccttcg ggcaaacaa agggcgcgcc ctgtatcgga cggcgatcg	27180	60
ggcgcgctgg ctgccggacg gcaacctcga gtttctcggt cgaaacgacg agcaggtgaa	27240	
ggtcccggt atccgcatcg agctggagga gatccgcgcg gcgttgctga agcaccggc	27300	
ggtcgctcaa gccgtggccg tgggtgcgca ggacgcgcg ggggacaagc ggctcgctgc	27360	
gtatgtctc ggacgcggag gacgcgcct gaccgcgcg gagctgcgc agtcctgtgag	27420	65
cgagcgattg cccgcgacca tgggtgcgct gtccttcgt gcgctcgacg ccctgcccct	27480	

cagccgaac ggcaaggtgg accgcccgc gctgccggag cccgagcgga gcgccggcgg 27540
 cgaggaccac gtcgcaccgc gcaacgccat cgaggaggag etcacacgaa tctgggcca 27600
 cgtacttggg gcaaagcggg tcggtgtgca cgacaatttc ttcgatctcg gcggccattc 27660
 5 cctgetgtct gtcgggtg atgatcggt cggccagcgg ttcgatcggc cgcctcgat 27720
 ggtegcacct ttacacctat cgaccgtggc gtcgctcgcg cggttccttg gcgaacgggc 27780
 gaacggcaag caatcgccga gggaggccgc ggcggacgtc acggagcgcg gccggcgccg 27840
 10 cctggaggcg cgggcgcggc gggcgaaggc catccgtggc ccgacctgac cggggcacc 27900
 ttccaagccc cgcggttct cgcacatccg cgcctcgag cgcgcgtcc agcgcgcgg 27960
 ttcgccagc aggaggcg agacgacggc ccaaggcctt cgtgggctct ttgccccga 28020
 atccggaagc tgcgcggcag ttcgtcgccc ctgcaatgct gccattgtag agctcctcg 28080
 15 ctcgccggc cctcttttct tgcggcccg cgcgattga cctcacatcc tgatcccttc 28140
 ttgctgctc cagaaagtga ttgacggcca gcgcgcgct tgagatcttc cggcgcgcg 28200
 cgatttcac gctccggcg gccgtgactg tcacctgca agggattata atgaaacata 28260
 acattgggtg gcttctaccc gccgcctcg cgacgcttg cttcgtcccg gcctgcagc 28320
 20 cgaatcacgg tgaggatg cgctccgtga cgtcagcaga gaggcgcgcg gcgcgagcg 28380
 ctgactgct cgcgctcggg gcgaagctcc aggcggcgct ggacggcgcc gccgcgcgc 28440
 aaaaggctcc gggagccgca gcgcgggtcc agagcgggga ctgtgtctgg cggggcgcca 28500
 25 cgggctctc ggacctggtc gcgagcacgc cgacgaagcc tggagatctc tttcggtcg 28560
 gcagcatcac caagaccttc gtctctacgc tgatactcat gctccgggca gaaggccggt 28620
 tgctgctcga cgacgcgggt tcgaagtatg tgaaggcat ccccgccggc gaccagatga 28680
 cgctgcgcca gatectcggt cacacgagcg ggtctctcga ttacacgtac agcccgcgcg 28740
 30 tcggccaaat gatcgagggt gatccgaccc gcgccttcgc gccggcagag ctcctcgccc 28800
 tcgccacggc cgaggcgccg tatttcgcgc cgggcgcggg ttttcgctat tcgaacacca 28860
 attacatcgt ggccggcctg gtggccgagg cgggtgtcgg cgggacgctc gccgggctgc 28920
 tccgaagcg eatectagac cctgtgggcc tgcgcacac gtatctggac ggcgccgagc 28980
 35 cgccggtcca agggctcacc cgcggctacg gcgactacgg cgcgggcttg gtcgacatca 29040
 ccgaccagct gtcgccacc gagcggtggg ccgcggcgcg cctggtgtcg aacgtcgatg 29100
 acctcaatcg cttctttgcc ctgctcatca gccacgagct gctctcgctg gacgagcttc 29160
 40 aggacatgac cacctggacc ccgacgatgt ggccccacga gcccgatat ggcccgcc 29220
 tcatcgagcg cgattctgct ctggctccc tcaacgggca ctgcggaatc atctggggt 29280
 ttcaatcgcc gtcgtacggg gtgcccggcc gcggcgacgc gatcccgcg ctcacaccc 29340
 ggagcgacgg cgacgcagcg cggtctgctg acgagctcgc gaaggctcgt aaagagcgt 29400
 45 gatcgaggcg gaatgggagc gcttcggcg gtggtgatgg cgcggcgcg tcagaacgcg 29460
 acgcgcagcc ccgcgctcag cgggctcgc cggggcgacg cggccacggc gcccgaccg 29520
 acgaggagcc gcgcgacggc gggcgcgctc ggcgcgctg ctcgcgcac ccgcgcttg 29580
 ccgaacacgt agagcggcag gccgacggcg accccggcca cccgcgcgag cgcggtggcg 29640
 50 atcgccacct cggacgcctc ggcgcgcgcg gcgctgctgt cgtcgtggct cgcgaagacc 29700
 agcaccgcgc cgctgaggat ggcgcgcgcc cccagggtcg tgaggacgag ccccgagatc 29760
 accatgacgc ggtgtgtcca ctcgctcgc cgtcctcga agtcgggaa cgcgcctc 29820
 gccgcgcga gctccagctc gatccggcg tgcctggcg gccgctcgt cgcggagccg 29880
 55 atctcgtgga cgcggcgggc ctggccctcc agcgcgcga tgcgcgctc gtgcgcggcg 29940
 ggggtctct cccacgtggc cctggcggg acccccgcca cggccggcg gacagaggcg 30000
 gccgacgcc gggctcgaggc gggcgccgc ggcggtctcg cggccacgga aggcgcgc 30060
 60 gccgcgggag gcgcggcgcg ctcgcggcc accgaaggcg ccgcgcgcg ggaaggcgcg 30120
 ggcgactcc cggccacgga aggcgcggcc gccgcgggag gcgcggcgcg ctcgctgca 30180
 acggagagcg cgggcgcgc gccagccagg cccagcgccc acgcgacgac acgacggcg 30240
 gcccaaccg cgcgcggcg cgcaaggcg aggtggacct gctccatcg cgcagcgtcg 30300
 65 cccctcgaca gggccgggtc aaggcgggg agtcgagag cagagacct ccgcgcgca 30360

ggaaacaggg gcgcggggcg cccgcgcggc ggctcgcgcg tcacccctcg cgcgggcggc 30420
 cgcgggcgcg cctccccctc cgggcgggccc gcgcgtcggc ggccacgcgg agcagctcct 30480
 ggaagtgcgg ctccaccggg cggaggtcga tgccgtccat gaacgacgtg aacgcgaagt 30540 5
 acggcagcag cgtctgccag gcgcgcagcc agcccgggag gtagcgcggc cgtccagca 30600
 cgcgggcggc gcgcagctcc gccagcaccg cgtatgtctc gagcgcgacg cggccgcaca 30660
 cgagcagctc cacctcgtcg cgcagcccg cgcgaaggac cgcggccagg aaggcgtaga 30720
 cctcgagcag cccggccagg tagcaggcgt ccttggtgaa cgcgcgcggc cctcgacga 30780 10
 gccgcgcgg gcacacgcgc tgggcgtcga agtaggcgtc gcggcgctcg gcgcgcgct 30840
 cgcgcagggtg cgggtacagg tcgaggaagc tcgcgccctg ctgggccatg tccacgagcc 30900
 gcacccgctc ggcgagccgg gtgaggcggc cgtggagag cgagcggctg tagagctcgg 30960 15
 cgaagatggc gagccccctc tgcgtgcgcg tgggtcgggg gccgcggcgg cgcaggaaag 31020
 cgcacgcgg ctgcgcgcg ccgttgtgcg cgggtgagcgc gtgcgtctcg acctcgtggt 31080
 gccacagccc ctgggcctcc cagcgcgcga aggtgcctc cggccggatg cggacccggc 31140
 tcatgcggcg gaccacctg gccgtgacgc gcgggtcgac ggtgatctcg aggtcgagcc 31200 20
 gcgggcgcgg gccggccag cgggcggcga gcatgtccg gagcgcgcg gcgtcgagcg 31260
 gctcctcctc gggatcgtg gcctcgtccc agcgtggac gcgcaggcgc tcgggtgaggt 31320
 gctcggcgag gtgatgtc ctgagcgagc cccgaagaa ccgcgagcgc gcgcgcgct 31380
 agagctcctg cgaccgcgcg gagaacgcgc ggggtgcgcg gccctcgagc agctccgcgg 31440 25
 cctggatctg cgcgcggacg ttgtcccga gccagcagc cgcgcgcgcg tcccgcgca 31500
 tggccccgag gagctcgcgc agctcggcga cgcgcgcgcg gaggcgcgct cgatcgacgc 31560
 ggtactcgac ctcggggagg cggctcctcg cgggcggcga gaagcgcctc tccacctcgc 31620
 gcggccaggg gatgtcctcg agcagcttga gggccttgc ctccgcagg cggccgcaca 31680 30
 cccgatcgag ctgctccagc acgggcgcgt cgtatgctat cgagcgcagg atcgccgaaa 31740
 ccgcgagacg ccggaacctg cattccctcg acgaggcagc gattgccatg ttccgtcgt 31800
 ttttggagcg ccgtcgtcgc gctcgcctgc gggctccggc gatccagcgc ggttgcatgc 31860 35
 agcaggggtg ttccggggct ggtcgcagag cgtcctttgg cccacacccg agacacgaat 31920
 gctccgcgcg gagcgcgggt gacctggac cgcgcggaga gccgatgata cggctccggc 31980
 gatgtcggag agttagctc aactcgaaga acaccgcgcg gcgtcaccg gadactgcta 32040
 cgggatgctg ggttcggtg tcgacgcga cgcgcgcgtc caggagacga tgggtcgcgc 32100 40
 ctggcgcgag ctggataagt tcgacggcg ctcgtcgtc cgcacctggc tgtaccgat 32160
 cgcgacgaac gtctgcctc acctgcgggc cgaccgcgcg cgcgcggcgc gcccacgca 32220
 ggaaggcccg gtcggcacgg tggacgacgc gctcgagacg cgcgcgcga cccactggct 32280
 ggagcccgct cccgacgcgc acgccctgcc ggcgacatc gacgcgcgg agcgggcgat 32340 45
 gctccgcag agcatccgc tcgcgttcgt cgcgcgcgtc cagcacctgc cgcgaagca 32400
 gcgcgcgcg ctgctgctca cggaggtgct cggctggtcc gccgcggagg tcgcgacag 32460
 cctcaacacc tcggtcgcg cgtcaacag cgcgtccag cgcgcgcgg cgacgtcgc 32520 50
 gagccgcgat ctggcgacg cgcgcctc gctgcggag ccgcagtcg cgtgctcga 32580
 ccgtacgtc aacgccttc agcggtaaga cgtcgacgc ctcacggcgc tgcgcacca 32640
 ggacgcgacc ctgctgatgc cgcgttcac cctgtggtc cgcggccacg agtcgatccg 32700
 cgcctggctc gtgggcccg gagcgggct cgcggggtc cggctcatcc cgacggcggc 32760 55
 gagggctcg cccgcgttc cgagtatcg cccgcgcgcg gagggcggcc accgggcctg 32820
 ggcgtcctc gtcctcgacg tcgcggggga ccgcctcgt agcatgacgt ccttctcga 32880
 caccgagacg ctcttccgc ggttcggcct gccgtcgat ctaccggcgt agcgcgggc 32940
 gccctgctc cctcgcgcg ggtgccctgc ctgcctagcc gcgggcgcgc ggctggcca 33000 60
 cgggcgcgcg gcttgccac ggggcgcgcg ccagcgacgg ggcgacgatt ttttctgag 33060
 cgaccgatga gtcctgacg gccgggggt ctacgggggt gaatccaaca cggaggcacc 33120
 catgaccgtg accatcgcca gcatcgatc tcgtgaccg gacctcatga cggggccca 33180
 ggccaaggcg ccggccgcg cgggcgcgcg cgcgcgcgc cgtccaggc gagcgtgtg 33240 65

ggcgggcgcg gtcctgagcg ggctggccac gctgttcctg acgttcgacg cgcgggtgaa 33300
 ggtgctgaag ctgttccccg cggaggcgte gaccgcccag ctcgggttcc cggcgcacct 33360
 cgtccccacc ctcggctacc tccagatcgc ttgcctcgtg gcctacctga tcccgcgac 33420
 5 cgcgggtgctc ggcgcgatcc tgtggaccgg ctacctgggc ggcgcgatcg cgatccacgt 33480
 gcggttcgag aacccgctct tcagccacac gctcttcccc atctacgtcg cgcggttcc 33540
 ctggggcggg ctctggctgc gcgaccgccc cgtgcgcgcg ctgaccgcga gcccgctgc 33600
 10 gcagggccga tgagcttcac gtttcacgag agtccatcac ggtaaaagga gaagcgagcc 33660
 atgaccacaa agaacccccg caagctcttc gtaacctgt cgcctccgga cctgaagcga 33720
 togatggagt tcttcagcaa gctcgggttc gaggttcaacc cgcagttcac ggacgagaag 33780
 gccgcctgca tggctgctag cgaggaggcc tatgtcatgc tctcgtgga gtcgttcttc 33840
 15 aagacgttca tgaagaagga gatctgcagc acgagcacgc acacggaagg gctcttcgcg 33900
 ctctcgtgca gcagccgggc cgaggtcgac gacatgggtga agaaggcggt cgcggcgggc 33960
 gggctgcacg cgatggatcc gcaggatcac ggcttcatgt acgggtggag cttctacgac 34020
 gtggatggcc accactggga ggtcatgtgg atggatccca agcgatcca gccgtagccg 34080
 20 acggggctgg gcgcgcgcgc tggagagacc cccgtgaggc ggggaggcgg gaggatcacc 34140
 gtcttcgtag cccacagcga tgcagtatcc gtcgcgcttc gtatcgaagc acgctgttta 34200
 cggcgcgctc agagcgctgc gcaggtgatg ccgagccgca gcagcgacac gggcacgagc 34260
 25 gtggctccga tggagatgag ccgagtcctc ccatgggtct cggggctcatg aatggatgag 34320
 taggggaact gctcctctgt cactcgtgc tcgacggcga cggcgaggcc gagctcgaag 34380
 tgcacggggc ctggaccgaa gatccagctc gcccccgcgc gagccccgac gaaaagcgtg 34440
 tcgcgctcga cgcacgggcc gtcgtccag cggggcgatc ccaccgcggt gtaggtgtgt 34500
 30 tccccgaagg aacccgcgag cgagagtcga agtcgcagcc gcgctcgcca cgcgacgccc 34560
 gctgtcgcgc cgacgcgcgc gaagctctcc ccgaaaggct tatccctgt ctcgatgaag 34620
 ccaccacact cgatgacgct gatcggttac gtgagcgcca gattgaggtg caccaccgcg 34680
 ctgtccgagc ccgagtagag gccggcgccc acctggaagg tgaatccat gtcggcgcg 34740
 35 gatecgcgcg caggagcgac gccaggggcg ctgccctcct gcgcgcgggc cgtcccgcag 34800
 caaagaaaga gggctgtcgc gaagaatcca agcgagatcg atcgaagtga gcgcatgtcg 34860
 ggccctggag catcgcgtgt accaggtgcg tcgtatctat gcggcgcgcc gccggcgccc 34920
 gccgcgctgg cctgtccgac gcgagatcac gaatccgcca tcgtccctcc gggccgcgcg 34980
 40 cgcctctggt tcgcctgcgg gcgtgcgcgc gcgctcgtgt ggcccatggc aacctgtctg 35040
 cgggtgtcgt cgaacagcac agagagtatc gcgtccgcaa caaccgcgcg acccgcgag 35100
 acgctcgtgg ggccccctgc ctccccactt catcataacg ccatcaggag cactcgacat 35160
 45 ttcatttctt cacctccact ggctgagggc gacggtgctc gtcacggcc ggttgcctcg 35220
 gcggttgcct tggcgggggt tctgacgccc ggaactaacg cttegagcgc tcccccttgc 35280
 tctcccgctt cttcagctcc tccagcaggt cgtcgaggcg ctcgtagctg cttcccgaga 35340
 agcgggcgta gttgtcgagc cagccgctgg cgtcctcgag cggcttggcc tcgacccgac 35400
 50 aaggcctccg ctgcgcgtcg cgcccgcgcg agatcaggcc cgcctcgtcc agcaccttga 35460
 ggtgcttggg gatcgcgggc tggctcatcg cgaacggctt cgccagctcg gtcaccgacg 35520
 cctcgccgga cgcgaggcgc gcgaggatcg ctgcgcgtgt cggatcgggc agcgacgca 35580
 acgttgcgtc gagggcgctc gacgggggtca ttgcataact ccttgggtata aaaaccagtt 35640
 55 agttatacaa cctgggggccc gggcggtcaa gcctccaggc gatggcggtt cggcccgggg 35700
 gctccgctcg cggcaccgcg gccgcgcggc tacgtgcgcg gcgcggtgag cactcctgc 35760
 agcgtggcgc cgaccacggg cttggtcagg tgcaggctga agccggccc cctggacctg 35820
 60 gcctgatcgt cgggcccgcg gtaccccgag agcgcacca ggtagagcgc ttgcgcgcg 35880
 ggcgcgggcc gcgcccggcg cgcgacctca taaccgtcga tgcggggcaa gccgatgtcc 35940
 acgaaggcca cctcgggggc cagctccaga agcttcttca cgccctccag cccgtccacc 36000
 gccaccgtca cctcgtgccc cagcgcctcg atgtacgccc gcatcaccgc gcgcacgtcc 36060
 65 tcccgctcct ccacgacgag cacccggcgc cggtcagcgc ccgcctcggg cgcctcggcg 36120

cgetgcgcgcg gaggcggcgcg cggctcgtcg cgetgcgcgcg gaggcggccc ctcgcgcgcgc 36180	
ggggggcggcc cggcgcgcgcg ggcaggetgc ggcgcgcgcgc cggggccgag cggcagggcgc 36240	
acgggtgaact cgetgcgcgcg gcccgccgcg ggcgcgcgcgc cggccacgcg gccgcgcgcgc 36300	5
agttccagga gccgcgcgcac cagcgtgagc ccgagcccca gcccgccgcg gtcgcgcgcgc 36360	
atggtctggt cgacctgcgt gaacagatcg aacaccttct cgagcatcgc cgcgggagtc 36420	
ccgcggccgcg tgcgcgcgcac ccgcagcagc gcctcgggcg cgcgcgcgcgc cgcgcgcgcgc 36480	
gtgaggcgca ccgagatcga gcccccggc ggggtgtact tcgcggcggt ggtcaggagg 36540	10
ttcgtcacca cctgctccag ccgcgcgcgc tcggcccgca tgcgaagtc cccgggcccc 36600	
accgacagcg acacgtcatg gcgcggggcc tcgacggcgc gcctcaccgc ggcggcggcg 36660	
ctctgcacca ccgcgcgcgag atcgacgtcc tcgaggcgca gctccaccgc gcccgcgctg 36720	15
atgcgcgaca cgtcgagcag atcgtcgacc agccgcacga ggtggcccat ctgcgcgcgc 36780	
gcgatctccc ggtagcgcgc cgacgcgggc cgtcgcgcgc cgcgcgcgcgc gacgagcgc 36840	
agcgacaggc tgatcgagc catcggttc cggagctcgc gcgcgagcat cgcgaggaac 36900	
tcgtccttgc gctgatcggc gagcttcagc goctcgacga gcgcctccac gcgcctccgc 36960	20
gcgcgcacct ggtcggtcac gtgaaacgc aacacgaaga cgcctcgac cgcgcgcgcgc 37020	
cgatcgcgca tcggctggtg gacgaagttg aagaacacct cctcgcgcgc gccgcgcgcgc 37080	
cggcgatcga gccgcaccgc gagctccttg ccgacgatgg gtcgcgcgcgc gccgaccacc 37140	
gcgtcgagga gctcccagat gccctgtccc tcgagctcgc ggaggcgggc ccggatgggc 37200	25
tcgcccacga gcgatcgacc gccgacgagc cgtcggtaga gccgggtgac cactcgaag 37260	
acgtgctccg gcccgcgag gatggcgatg gcccccgggg cctgcatgaa gaggtcgttc 37320	
aggtactggc gctgccccgc gccctcgcgc cgcggcgcg cgagctcgac gtggatgcgc 37380	30
acccgcgcga ggagctcctt cgcgggagaa ggttcacga ggaagtcgc gccgcgcgcgc 37440	
tcgaggctgt cgaacgcgcgc cctctgcgc gccgcgcgcgc agagcatcac cagcgcgagc 37500	
ccgcgggtgc gatcgtcggc gccgagcgc ctgagcaggc cgaagccgc gagccgcgcgc 37560	
atcatcacgt cggtgagcgc gagatccgc ggggtgggcgc gggcgcgctc caggcgcgcc 37620	35
cgaccgctgc ccaaggcctc caccgtccac cctccgcca cgagcagccg cagcgcgta 37680	
tcgcgcgtgt ccgcggttgc gtcggcgagc aggaacgcgc ccggcagcct cccggcggcgc 37740	
ccctgcgcgc ccggcgcgga cgcggcgcc tgcgcgcgc gcgacactg cgcggcctgc 37800	
tcgaggaagg gcgcggcgtc cgcgcgcgc gccgcgcgc ccgaggccgc cgcgcgc 37856	40

Patentansprüche

1. DNA-Sequenz, deren Expressionsprodukte die enzymatische Biosynthese, Mutasynthese oder Partialsynthese von Polyketid- oder Heteropolyketidverbindungen bewirken oder daran beteiligt sind. 45
2. DNA-Sequenz nach Anspruch 1, wobei es sich bei den Polyketid- oder Heteropolyketidverbindungen um Epotilone handelt.
3. DNA-Sequenz nach einem der vorhergehenden Ansprüche, wobei die DNA-Sequenz Regulationselemente (ORF9, ORF11 und ORF12) und Transkriptionsregulatoren (ORF10, ORF13 und ORF14) aufweist und die Expressionsprodukte eine tRNA-Synthetase (ORF1), Monooxygenase (ORF2), Aminotransferase (ORF3), Tyrosin/DOPA-Decarboxylase (ORF4), 3-Oxoacyl-ACP-Reduktase (ORF5), Polyketidsynthase (ORF6), Peptidsynthetase (ORF7) und Transpeptidase (ORF8) umfassen. 50
4. DNA-Sequenz nach einem der vorhergehenden Ansprüche, wobei die DNA aus Myxobakterien stammt.
5. DNA-Sequenz nach einem der vorhergehenden Ansprüche, wobei die DNA aus Sorangium-Stämmen stammt. 55
6. DNA-Sequenz nach einem der vorhergehenden Ansprüche, wobei die DNA aus Sorangium cellulosum stammt.
7. DNA-Sequenz nach einem der vorhergehenden Ansprüche, wobei die DNA ausgewählt ist unter:
 - (a) der folgenden DNA-Sequenz: oder deren komplementärem Strang,
 - (b) DNA-Sequenzen, die unter stringenten Bedingungen an die Proteine kodierenden Regionen der in (a) definierten DNA-Sequenzen oder an Fragmente davon hybridisieren,
 - (c) DNA-Sequenzen, die wegen der Degeneration des genetischen Kodes an die unter (a) und (b) definierten DNA-Sequenzen hybridisieren,
 - (d) alle Variationen und durch Substitution, Insertion oder Deletion von Nucleotiden entstandene Mutanten der unter (a) bis (c) definierten DNA-Sequenzen, die isofunktionelle Expressionsprodukte ergeben. 60
8. Rekombinierter Expressionsvektor, der eine DNA-Sequenz nach einem der Ansprüche 1-7 enthält. 65
9. Prokaryotische oder eukaryotische Zelle, die mit einer DNA-Sequenz nach einem der Ansprüche 1 bis 7 oder mit einem rekombinierten Expressionsvektor nach Anspruch 8 transformiert oder transfiziert ist.
10. Zelle nach Anspruch 9, wobei die Zelle von Myxobakterien stammt.

11. Zelle nach Anspruch 9, wobei die Zelle von einem Sorangium-Stamm stammt.
12. Zelle nach Anspruch 9, wobei die Zelle von Sorangium cellulosum stammt.
13. Zelle nach Anspruch 9, wobei die Zelle von einem Streptomyces-Stamm stammt.
14. Zelle nach Anspruch 9, wobei die Zelle von Escherichia coli stammt.
5 15. Verfahren zur enzymatischen Biosynthese, Mutasynthese oder Partialsynthese von Polyketid- oder Heteropolyketidverbindungen, bei dem eine Zelle nach einem der Ansprüche 9 bis 14 in einem geeigneten Kulturmedium kultiviert und die Polyketid- oder Heteropolyketidverbindung aus dem Medium isoliert wird.
16. Verfahren nach Anspruch 15, wobei die Polyketid- oder Heteropolyketidverbindung ein Epothilon ist.

Hierzu 2 Seite(n) Zeichnungen

10

15

20

25

30

35

40

45

50

55

60

65

Fig. 1

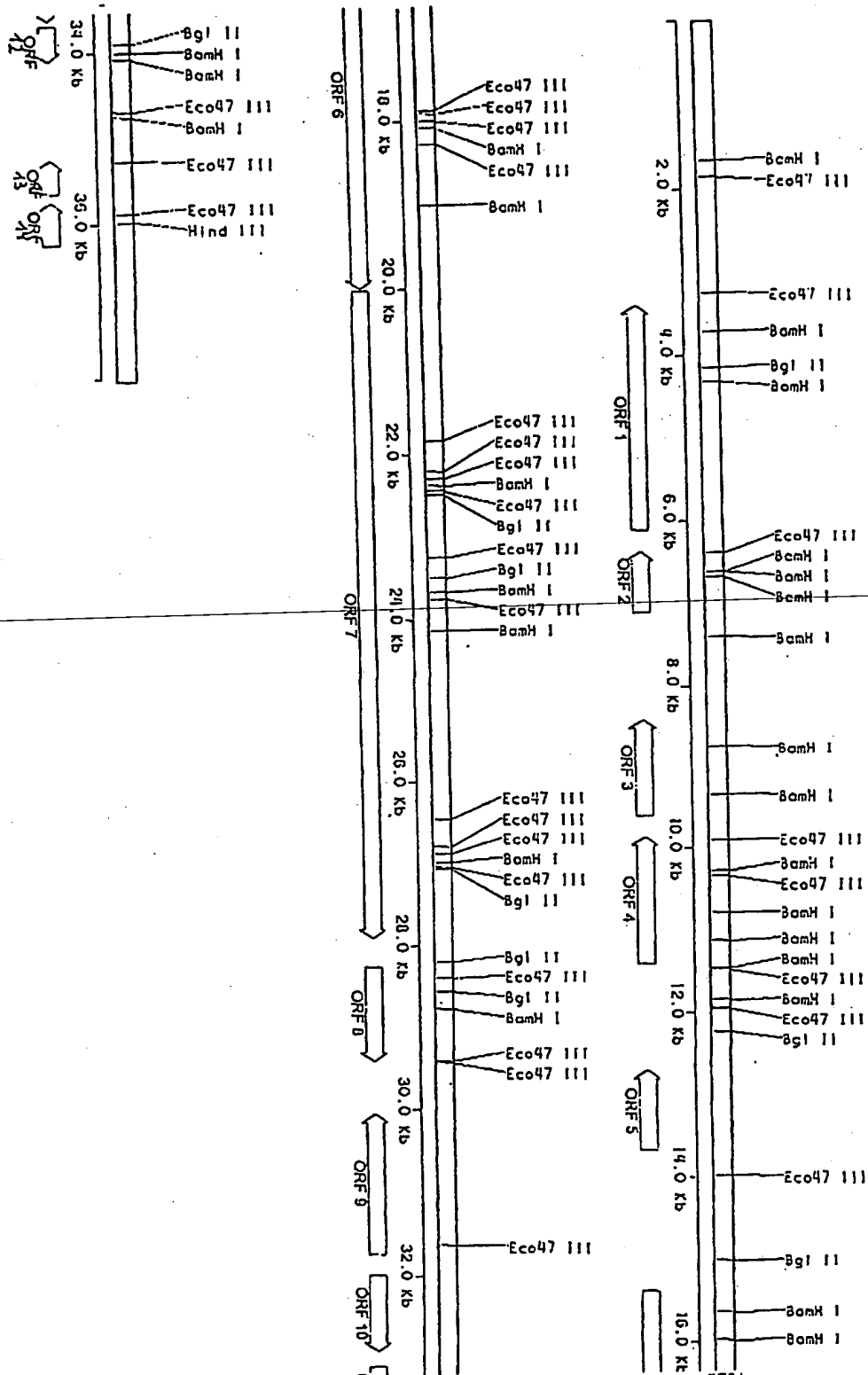


Fig. 2

	Gen/Funktion	Position
ORF 1	RNA-Synthetase	6100-3397
ORF 2	Monooxygenase	7111-6374
ORF 3	Aminotransferase	9550-8433
ORF 4	Tyrosin/DOPA-Decarboxylase	11393-9854
ORF 5	3-Oxoacyl-ACP-Reduktase	12212-13656
ORF 6	Polyketidsynthase	15374-19984
ORF 7	Peptidsynthetase	20003-27889
ORF 8	Transpeptidase	28251-29402
ORF 9	Regulationsselement	31720-30040
ORF 10	Transkriptionsregulator	31982-32932
ORF 11	Regulationsselement	33128-33613
ORF 12	Regulationsselement	33661-34076
ORF 13	Transkriptionsregulator	35616-35255
ORF 14	Transkriptionsregulator	36242-35730